OPTIMIZATION FOR BURSTING NEURAL MODELS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Joseph Hua Tien

January 2007

OPTIMIZATION FOR BURSTING NEURAL MODELS

Joseph Hua Tien, Ph.D.

Cornell University 2007

This thesis concerns parameter estimation for bursting neural models. Parameter estimation for differential equations is a difficult task due to complicated objective function landscapes and numerical challenges. These difficulties are particularly salient in bursting models and other multiple time scale systems. Here we make use of the geometry underlying bursting by introducing defining equations for burst initiation and termination. Fitting the timing of these burst events simplifies objective function landscapes considerably. We combine this with automatic differentiation to accurately compute gradients for these burst events, and implement these features using standard unconstrained optimization algorithms. We use trajectories from a minimal spiking model and the Hindmarsh-Rose equations as test problems, and bursting respiratory neurons in the preBotzinger complex as an application. These geometrical ideas and numerical improvements significantly enhance algorithm performance. Excellent fits are obtained to the preBotzinger data both in control conditions and when the neuromodulator norepinephrine is added. The results suggest different possible neuromodulatory mechanisms, and help analyze the roles of different currents in shaping burst duration and period.

BIOGRAPHICAL SKETCH

Joseph Tien attended high school at Hawken Upper School in Gates Mills, Ohio. He studied Ecology and Evolutionary Biology at Princeton University for his bachelor's degree, and Applied Mathematics at Cornell University for his doctorate.

To my family

ACKNOWLEDGEMENTS

I am very grateful to my advisor, John Guckenheimer, for consistent guidance and support throughout. Thanks also to Ron Harris-Warrick and Lars Wahlbin for serving on my committee.

My path through graduate school has been a wandering one. Several individuals took a chance and gave me the opportunity to study applied math. John Guckenheimer and Steve Strogatz encouraged me to explore my interests and apply to CAM. Carlos Castillo-Chavez allowed a curious ecology and evolutionary biology graduate student to attend his computational biology training grant meetings. To them I am deeply thankful.

I had the good fortune to work with Mary Lou Zeeman through the IGERT program. The National Science Foundation provided funding for several years of my graduate studies. Thanks to Dolores Pendell, without whom CAM would swiftly crumble. CAM has been home for these past years – may it always keep its good cheer.

Thanks most of all to my parents and brother for their love through the years.

TABLE OF CONTENTS

1 Introduction **1**
- 1.1 Fast-slow systems . . . 3
- 1.2 Bursting and fast-slow dissection . . . 5
- 1.3 Optimization . . . 6
- 1.4 Automatic differentiation . . . 8

2 Algorithms **10**
- 2.1 Overview . . . 10
- 2.2 Computing solution trajectories . . . 12
  - 2.2.1 Bursting event defining equations . . . 12
  - 2.2.2 IVPs without events . . . 13
  - 2.2.3 IVPs with events . . . 16
  - 2.2.4 Periodic orbits without events . . . 17
  - 2.2.5 Periodic orbits with events . . . 18
- 2.3 Objective functions and their gradients . . . 18
  - 2.3.1 IVPs without events . . . 20
  - 2.3.2 IVPs with events . . . 20
  - 2.3.3 Periodic orbits without events . . . 23
  - 2.3.4 Periodic orbits with events . . . 24
- 2.4 Taking the optimization step . . . 26
  - 2.4.1 Steps for a bursting periodic orbit . . . 27

3 Test problems **30**
- 3.1 A minimal spiking model . . . 30
  - 3.1.1 Motivation and bifurcations of the minimal model . . . 31
  - 3.1.2 Optimization and the minimal model . . . 34
- 3.2 A bursting example: Hindmarsh-Rose . . . 36
  - 3.2.1 Hindmarsh-Rose equations . . . 37
  - 3.2.2 Continuation and canards . . . 38
  - 3.2.3 Optimization algorithm performance: fitting a transient burst 43
  - 3.2.4 Periodic bursters: fitting period and burst duration . . . 49

4 PreBotzinger neurons **54**
- 4.1 Biology of the preBotzinger complex . . . 55
  - 4.1.1 Ion channels in the preBotzinger complex . . . 56
  - 4.1.2 Modulation of the respiratory rhythm . . . 58
- 4.2 The model of Butera, Rinzel and Smith . . . 60
- 4.3 Optimization for Butera et al's model . . . 62
  - 4.3.1 Control data fits . . . 64
  - 4.3.2 Norepinephrine data fits . . . 69
- 4.4 Discussion . . . 74

5 Conclusions and future work **80**
5.1 Future directions 81

LIST OF TABLES

3.1 Optimization results for transient bursts in Hindmarsh-Rose. . . . 50

4.1 Phase space variables for Butera et al's model. 62
4.2 Parameters for Butera et al's model. 63
4.3 Empirical burst data for Cd^{2+}-insensitive pacemakers. 64
4.4 Initial fits to preBotzinger control data. 65
4.5 Parameter sensitivities for the Set 1 burster. 67
4.6 Obtained fits for Cd^{2+}-insensitive pacemakers in the control setting. 70
4.7 Parameter sensitivities for the Control 1 burster. 72
4.8 Obtained fits for Cd^{2+}-insensitive pacemakers with norepinephrine. 73

LIST OF FIGURES

2.1 Burst event surfaces. 14
2.2 Local vs. global burst event checking. 15

3.1 Bifurcations in the minimal model. 33
3.2 Objective function landscape for the minimal model. 35
3.3 Optimization steps and contours of the minimal model landscape. . 35
3.4 Burst orbit in the Hindmarsh-Rose system. 39
3.5 Continuation curve of fixed period bursters for Hindmarsh-Rose. . 40
3.6 Canards in the Hindmarsh-Rose equations. 41
3.7 Perturbations from canard orbits in the Hindmarsh-Rose system. . 42
3.8 Voltage time series landscape for Hindmarsh-Rose. 45
3.9 Event timing landscape for Hindmarsh-Rose. 46
3.10 Spike number and objective function plots for Hindmarsh-Rose. . . 47
3.11 Fitting burst duration by gradient descent and continuation of fixed period orbits. 52

4.1 Four bursting periodic orbits for the preBotzinger model. 66
4.2 Optimization algorithm steps for fitting the preBotzinger model to data. 74
4.3 Continuation of fixed period bursters for preBotzinger neurons. . . 78

LIST OF SYMBOLS

f	right hand side of a differential equation
λ	parameters
Φ	flow map of the vector field
F	multiple shooting map
$\mathbf{\Delta}$	trajectory points which are free to vary
G	objective function to be minimized
Γ	event defining function
M_0	critical manifold (collection of fast subsystem fixed points)

Chapter 1
Introduction

Mathematical neuroscience has been a rich area of collaboration between mathematicians and neurobiologists ever since Hodgkin and Huxley's [59] Nobel Prize-winning work on the squid giant axon. Similar ordinary differential equation (ODE) models have since been used to model a variety of excitable cell types exhibiting an array of electrical activity patterns, including quiescence, tonic spiking, and bursting. Development and calibration of these models is an active research area in neuroscience and mathematical neuroscience. Bursting trajectories, where a cell alternates between quiescent and active phases, are of particular interest. Bursting cells occur throughout the body and are associated with biologically important events such as hormone secretion and muscle contraction [68]. Mathematically, bursting models can be quite complicated due to the presence of multiple time scales in the equations. Automated methods for fitting these bursting models to data would greatly aid in these bursting studies, but little work has been done in this area [99]. In general, algorithm development for ODE parameter estimation is incomplete. This is likely due to some particular challenges of optimization for ODEs. Objective function landscapes for ODEs can be complicated, possessing many local minima [27, 31, 109]. Solution trajectories can change drastically in response to a small change in parameters, for example when the system passes through a bifurcation. There are also numerical challenges for optimization for ODEs. One obstacle is the accurate computation of derivatives of solution trajectories, which are needed for gradient-based optimization algorithms. Closed form solutions are typically unavailable for ODEs, and thus analytical expressions for these derivatives are usually unavailable as well. Additional numerical challenges

arise in the context of ODEs possessing multiple time scales. These ODEs can possess certain solution trajectories which are extremely sensitive to parameters, making optimization difficult.

This thesis concerns smooth local optimization methods for bursting neural models. There are two main aspects to the work. The first is the use of automatic differentiation and Taylor series integration [50] to accurately compute derivatives of solution trajectories. The second is defining different objective functions which generate less complicated objective function landscapes. One way this can be accomplished is to restrict the optimization algorithm to a particular subset of solution trajectories to the ODE, such as periodic solutions. Casey [15] worked on both of these aspects by using automatic differentiation for optimizing (non-burst) periodic orbits. We build upon Casey's work and define objective functions which exploit the geometry underlying bursting. Specifically, we introduce defining equations for burst initiation and termination. These defining equations can be used together with automatic differentiation to compute the parameter sensitivities of biological quantities of interest such as burst timing and duration, period, and duty cycle. These sensitivities can be used in gradient-based optimization methods for fitting models to data. They can also help analyze the role of different ion currents in shaping burst characteristics, and suggest possible neuromodulatory targets.

Several simplified, non-mechanistic neural models have been proposed which retain the qualitative firing patterns seen in Hodgkin-Huxley style models [35, 58]. We consider both mechanistic and qualitative models in this thesis. We test our algorithms using the Hindmarsh-Rose equations, a polynomial system that can exhibit bursting [58]. We also draw additional test problems from a minimal spiking

model possessing a single fast sodium current and a slower potassium current, with all other currents much slower than the potassium current. Finally we examine Butera et al' s [11] model of preBotzinger complex respiratory neurons, and use our algorithms to fit this model to measured burst characteristics under different experimental conditions [125].

The remainder of this thesis is organized as follows. Chapter 1 discusses background material for fast-slow systems, neural models and bursting, and optimization. Chapter 2 discusses the algorithms which we develop and use. This includes methods for computing different types of solution trajectories (e.g. general initial value problems (IVPs), burst periodic orbits) and their derivatives. We apply our algorithms to test problems drawn from a minimal spiking model [51] and the Hindmarsh-Rose equations [58] in chapter 3. The examples show the utility of automatic differentiation and the burst event defining equations, and also illustrate some of the features of burst orbits that make optimization challenging. Chapter 4 examines a model of respiratory cells in the preBotzinger complex. We use our algorithms to fit Butera et al's model [11] to experimental measurements of burst characteristics both in control settings and in the presence of the neuromodulator norepinephrine [125]. The optimization algorithms and sensitivity calculations suggest roles for the different constituent currents in shaping bursts. Chapter 5 finishes with some concluding remarks and future directions.

1.1 Fast-slow systems

The material here follows Jones [74] and uses material from Tien et al [120].

Fast-slow systems can be written in the canonical form

$$\begin{aligned} \dot{x} &= f(x, y, \epsilon) \\ \dot{y} &= \epsilon g(x, y, \epsilon), \end{aligned} \tag{1.1}$$

where $x \in \mathbb{R}^m, y \in \mathbb{R}^n$, and ϵ is a small positive parameter. Rescaling time allows us to rewrite system (1.1) as

$$\begin{aligned} \epsilon x' &= f(x, y, \epsilon) \\ y' &= g(x, y, \epsilon), \end{aligned} \tag{1.2}$$

where the $'$ indicates differentiation with respect to rescaled time. System (1.1) is called the fast system and system (1.2) the slow system. Letting $\epsilon \to 0$ in system (1.1), we obtain an m-dimensional system with parameters y:

$$\begin{aligned} \dot{x} &= f(x, y, 0) \\ \dot{y} &= 0. \end{aligned} \tag{1.3}$$

In system (1.2), the limit $\epsilon \to 0$ only makes sense if $f(x, y, 0) = 0$. The solution set $f(x, y, 0) = 0$ specifies in general an $n-$dimensional manifold, given (locally) by solving $f(x, y, 0) = 0$ for x in terms of y. This is called the *critical manifold*.

Let M_0 denote the critical manifold. M_0 is said to be *normally hyperbolic* if fixed points of

$$\dot{x} = f(x, y, 0)$$

are hyperbolic, where y is regarded as a parameter. Given a normally hyperbolic critical manifold, the Fenichel theorems [34] state that for positive ϵ sufficiently small, there exists a locally invariant manifold M_ϵ of $O(\epsilon)$ distance away from M_0.

M_ϵ possesses stable and unstable manifolds of the same dimensions as the stable and unstable manifolds of M_0.

The van der Pol system is a classical example of a fast-slow system [123]. The flow consists of trajectory segments which closely hug stable components of the critical manifold, together with fast "jumps" where the system moves from one stable component of M_0 to another. This behavior of spending the majority of time near attractors of the fast subsystem together with fast transitions between attractors is the same pattern seen in bursting systems.

1.2 Bursting and fast-slow dissection

Bursting is an electrical activity pattern characterized by alternation between quiescent and active phases, with action potentials occurring in the active phase [68]. This is an ubiquitous phenomenon observed in cells in the central nervous system [22, 23, 45, 69, 128], the endocrine system [16, 111, 115], and the cardiac system [1]. Bursting is associated with muscle contraction for locomotion [28, 45, 69, 126], respiration [33], and digestion [53], can trigger secretory activity [105, 111, 115], and is thought to play a role in learning [95]. Bursting cells often play pacemaker roles [101] and have been studied extensively in the context of central pattern generators [28, 53, 97, 126]. There are many medical pathologies which motivate study of bursting cells, including diabetes, Parkinson's disease [94], cardiac arrhythmias, sudden infant death syndrome, and epileptic seizures [2].

Mathematically, bursting occurs in fast-slow systems when there is alternation between close tracking of a (hyperpolarized) fixed point with following a family of stable periodic orbits in the fast subsystem. Tracking hyperpolarized fixed points corresponds to quiescence, and following a family of periodic orbits corresponds to

spiking. The pioneering work of Rinzel and Lee [104] on fast-slow "dissection" of an *Aplysia* R15 neuron model helped illuminate these bursting mechanisms. Transitions between the quiescent and active phases are associated with bifurcations in the fast subsystem. Classification schemes based upon the types of bifurcations that occur have been proposed [64].

Parts of this thesis concern characterizing burst initiation and termination events. The key points are that burst initiation is associated with firing an action potential, and termination is associated with close passage to a stable component of the critical manifold.

1.3 Optimization

Optimization is concerned with minimizing an *objective function*. In the context of parameter estimation, the objective function is a measure of goodness of fit between model to data, with the minimization occurring over some set in parameter space. Optimization is a large branch of mathematics and there are many different algorithms to choose from [24, 36]. Algorithms may be unconstrained or constrained, and may use derivatives or be gradient-free. Within each of these categories there are many subcategories. For example, constrained methods may be solved by interior or exterior point methods, sequential quadratic programming, or penalty functions [36]. Gradient-free methods include the Nelder-Mead simplex method [91], simulated annealing [72], and genetic algorithms [88].

This thesis largely concerns smooth, local optimization algorithms. To illustrate the basic ideas we discuss trust-region methods. This includes the Levenberg-Marquardt algorithm [78, 83, 89], which will be used in many of the examples we consider.

Let $\lambda \in \mathbb{R}^p$ be parameters and $G : \mathbb{R}^p \to \mathbb{R}$ our objective function. An optimization algorithm generates a sequence of parameter values which hopefully converge to a minimum of G. Let λ_c denote the current parameter value. Trust region methods find the next parameter value by forming a local model of G, and then minimizing this local model within some region where the model is "trusted" [24]. A quadratic local model is often used:

$$m_c(s) = G(\lambda_c) + s^T \nabla G_c + \frac{1}{2} s^T H_c s. \tag{1.4}$$

When the Hessian H_c is positive definite the local model is minimized by the Newton step:

$$s_N = -H_c^{-1} \nabla G_c. \tag{1.5}$$

This Newton step is modified if s_N goes beyond the trust region. There are various ways to modify the step to stay within the trust region. One of these is the hook step:

$$s(\mu) = -(H_c + \mu I)^{-1} \nabla G_c. \tag{1.6}$$

The value of μ is chosen depending upon the trust radius. If the Newton step lies within the trust region, $\mu = 0$. Otherwise μ is chosen so that $s(\mu)$ lies on the boundary of the trust region. Note that as μ increases, the parameter step direction approaches the gradient step direction. Computing $s(\mu)$ gives a proposed parameter step. Whether this step is accepted or rejected is typically based upon the value of the objective function at $\lambda_+ = \lambda_c + s(\mu)$. That $G(\lambda_+) < G(\lambda_c)$ is a typical acceptance criterion. The size of the trust region is updated based upon the

match between the predicted and actual values of $G(\lambda_+)$, with expansion following a good match and contraction after a poor match.

In the context of data fitting a typical form for G is $G = \frac{1}{2}\|R(\lambda)\|_2^2$, where R is a residual vector. In this case $\nabla G = J^T R$, where J is the Jacobian of R. The entries of the Hessian are given by

$$H_{ij} = (J^T J)_{ij} + \sum_k \left(R_k \frac{\partial^2 R_k}{\partial \lambda_i \partial \lambda_j} \right). \tag{1.7}$$

Levenberg-Marquardt makes an affine approximation to the residual vector R. This gives

$$H \approx J^T J. \tag{1.8}$$

Computing (1.8) involves only first order derivatives and the expression has the added benefit of being positive definite.

This is a very brief discussion of optimization algorithms. Further details can be found in Dennis and Schnabel [24] and Fletcher [36].

1.4 Automatic differentiation

Automatic (also called algorithmic) differentiation (AD) is a method for computing derivatives to high accuracy without expression swell [47]. This is done by taking advantage of the fact that a computer evaluates a function as a composition of elementary operations whose derivatives are known. Evaluation of an expression can be represented as a tape which records the elementary operation and argument values involved at each step. Automatic differentiation steps along this tape, using the known derivative rules for each elementary operation together with the

numerical values of the arguments to propagate a number at each step. Thus automatic differentiation is based upon repeated use of the chain rule. Because AD propagates a numerical value rather than a symbolic expression at each step, AD does not suffer from the expression swell that occurs for symbolic differentiation. Like symbolic differentiation, however, the derivatives are calculated very accurately, essentially to the same accuracy as the original function. This is because the derivatives of the elementary operations involved in evaluating the function are known exactly.

Automatic differentiation's Taylor coefficient and adjoint modes are particularly useful for differential equations. Given the right hand side of an ODE, AD can compute the Taylor coefficients of the flow of the system through a bootstrapping technique. This generates a degree d approximation to the flow: $\Phi(t, x_0) \approx \Phi_d(t, x_0)$. Adjoint mode can be used to compute the Jacobian of the flow $D_x\Phi_d$ [47, 48, 96].

Guckenheimer and Meloon [50] used the automatic differentiation package ADOL-C [48] to implement a Taylor series integrator and periodic orbit solver based upon these features. This included modifying ADOL-C to allow computation of derivatives of the flow with respect to parameters. Phipps [96] developed ADMC++, an automatic differentiation package in Matlab including these features. ADMC++ also includes solvers for hybrid systems. Casey [15] utilized ADMC++ in developing optimization algorithms for periodic orbits. We primarily use ADMC++ in this thesis, but also use ADOL-C in places.

Chapter 2

Algorithms

2.1 Overview

The approach we take to optimization of ODEs is a simple one involving finding successive solutions to the ODE. Given a current parameter value, we find a solution trajectory. From this solution trajectory we compute the objective function G. Gradient-based methods also require the gradient ∇G and possibly higher derivatives of G. We then determine a parameter step based upon this information, after which this process is repeated.

The type of solution trajectory to find depends upon the problem type. General initial value problems (IVPs) are treated differently than periodic orbits. For both IVPs and periodic orbits we may have *event detection* in order to locate special points along solution trajectories. For example, we may wish to locate burst initiation and termination events which mark the start and stop of the active phase of a burst. This gives a total of four different problem types: IVPs with and without event detection, and periodic orbits with and without events.

Solution trajectories for each of these problem types are given by solving defining equations of the form

$$F(\mathbf{\Delta}(\lambda), \lambda) = 0. \tag{2.1}$$

For brevity we are abusing notation: the definitions of F and $\mathbf{\Delta}$ vary with the problem type and are given in the following sections. Conceptually, $\mathbf{\Delta}$ denotes the points on discrete curve representation of trajectories which are free to vary, and

F is a multiple shooting map induced by the flow of the vector field. These ideas are based upon the work by Guckenheimer and Meloon [50].

The objective function $G : \mathbb{R}^p \to R$ depends upon parameters $\lambda \in \mathbb{R}^p$ through the dependence of solution curves on parameters. Thus $G = G(\mathbf{\Delta}(\lambda))$, giving

$$D_\lambda G = D_{\mathbf{\Delta}} G \cdot D_\lambda \mathbf{\Delta}. \tag{2.2}$$

Computing $D_\lambda G$ involves computing $D_\lambda \mathbf{\Delta}$, which can be accomplished by differentiating (2.1) with respect to λ:

$$D_{\mathbf{\Delta}} F \cdot D_\lambda \mathbf{\Delta} + D_\lambda F = 0. \tag{2.3}$$

Provided $D_{\mathbf{\Delta}} F$ is square and nonsingular, we can compute $D_\lambda \mathbf{\Delta}$ by

$$D_\lambda \mathbf{\Delta} = -(D_{\mathbf{\Delta}} F)^{-1} D_\lambda F. \tag{2.4}$$

Section 2.3 discusses when this is the case.

The basic optimization algorithm is outlined in Algorithm 1. The key steps involved are computing solution trajectories of the appropriate type (i.e. solving $F = 0$), and computing the parameter sensitivities of these solutions (calculating $D_\lambda \mathbf{\Delta}$). We use automatic differentiation and Taylor series integration to compute solution trajectories [50, 96]. This allows high accuracy computation of $D_\lambda \mathbf{\Delta}$.

The remainder of this chapter is organized as follows. Section 2.2 discusses computation of solution trajectories for the different problem types. Definitions for $F, \mathbf{\Delta}$, and the burst defining equations are given here. Section 2.3 considers objective functions and their gradients. Calculating objective function gradients involves $D_\lambda \mathbf{\Delta}$, which can be obtained from (2.4) when $D_{\mathbf{\Delta}} F$ is invertible. Proofs

for when $D_{\mathbf{\Delta}}F$ is invertible are given in this section. Section 2.4 walks through one iteration of Algorithm 1 for a bursting periodic orbit.

2.2 Computing solution trajectories

To find solution trajectories (solutions to $F = 0$) we use the same multiple shooting framework used by Guckenheimer and Meloon [50]. What is different here is the use of events. Specifically, we identify burst event points which denote burst initiation and termination. These can occur both for non-periodic trajectories (e.g. transient bursts) and for periodic orbits.

2.2.1 Bursting event defining equations

A bursting system exists either in a *quiescent* or in an *active* state. We monitor for burst initiation events during the quiescent state, and for burst termination while the system is active. An initiation event results in a state transition from quiescent to active, while a termination event results in a transition from active to quiescent.

We say that burst initiation occurs when a voltage threshold is crossed:

$$\Gamma_{init}(z) = \pi_v(z) - v_{init}. \tag{2.5}$$

Here π_v denotes projection to the voltage coordinate and v_{init} is a voltage threshold specified by the user.

There are several different choices for defining burst termination. One possibility is to define burst termination as entry into a neighborhood of a stable component of the critical manifold. This gives

$$\Gamma_{term}(z) = \frac{1}{2}\|Sf(z)\|_2^2 - \delta, \tag{2.6}$$

together with the requirement that $\dot{\Gamma}_{term} < 0$. The size of the neighborhood is determined by δ, and S is a diagonal scaling matrix. We need an additional requirement to ensure that we are passing close to a stable component of M_0 and not, for example, close to a saddle component. One check is to take a proposed event point, locate a point on M_0 which minimizes the distance to the proposed event point, and then check the stability of the point on M_0. We call this *local* termination checking. An alternative check uses that the quiescent phase corresponds to a contiguous time period of at least $\mathcal{O}(1)$ of slow time during which no spiking occurs. Thus to check a proposed termination event, we integrate forwards in time for $\mathcal{O}(1)$ slow time and see if any initiation events occurred. We will call this *global* termination checking.

A schematic illustrating the burst initiation and termination defining equations is given in Figure 2.1. Figure 2.2 shows termination events detected using (a) local and (b) global event checking. Note that the two methods result in different event points. The state-dependent event detection discussed here is facilitated by ADMC++'s hybrid vector field class.

2.2.2 IVPs without events

Solving IVPs without events is the simplest situation. In this case the time mesh points are fixed and $\mathbf{\Delta} = (z_1, \ldots, z_m)$ consists of the phase space points only. The initial condition z_0 we will consider fixed. The multiple shooting map $F : \mathbb{R}^{m \cdot zdim} \to \mathbb{R}^{m \cdot zdim}$ is defined by

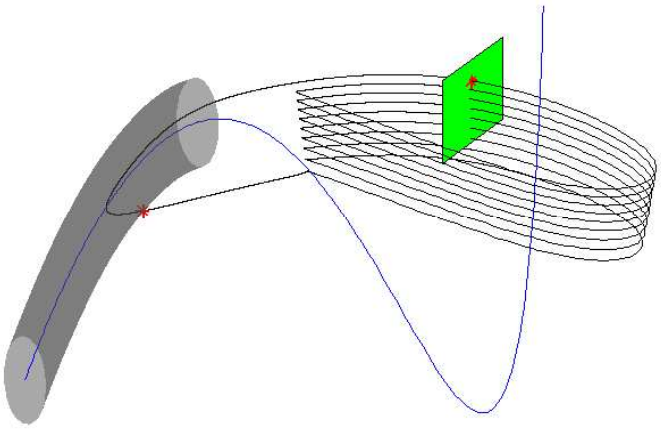
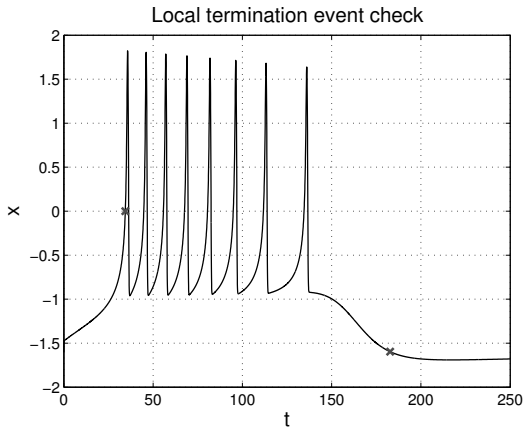
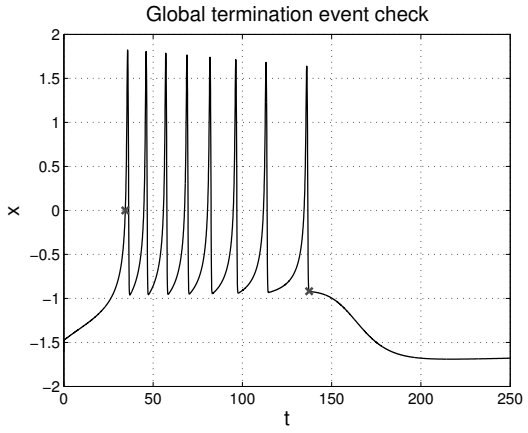


Figure 2.1: Schematic of a burst trajectory together with surfaces which define burst events. Burst initiation corresponds to crossing the green cross-section, and burst termination corresponds to entering the gray tube. M_0 is drawn in blue, the burst periodic orbit in black, and event points are in red.



(a)



(b)

Figure 2.2: Burst initiation and termination events detected for a transient burst in the Hindmarsh-Rose equations (3.5). Events are marked with an 'x'. (a) Local event checking. (b) Global event checking.

$$F_{z_0,\mathbf{t}}(z_1,\ldots,z_m) = \begin{pmatrix} \Phi(z_0, t_1 - t_0) - z_1 \\ \vdots \\ \Phi(z_{m-1}, t_m - t_{m-1}) - z_m \end{pmatrix}. \tag{2.7}$$

Solutions to $F = 0$ are obtained by numerically integrating the system. Integrating with ADMC++'s Taylor series integrator also gives the Jacobian of the flow between mesh points.

2.2.3 IVPs with events

When considering events we cannot use a fixed time mesh because the event times must be solved for. Let σ_i, $i = 1,\ldots,n_e$ denote event point indices, with n_e the number of events. For IVPs with events we thus have $\mathbf{\Delta} = (t_{\sigma_1},\ldots,t_{\sigma_{n_e}}, z_1,\ldots,z_m)$. Let $\mathbf{z} = z_1,\ldots,z_m$, and $\mathbf{t_{ev}} = t_{\sigma_1},\ldots,t_{\sigma_{n_e}}$. Let Γ_{σ_i} denote the ith event defining function, $i = 1,\ldots,n_e$. Then the definition of F is:

$$F(\mathbf{t_{ev}}, \mathbf{z}) = \begin{pmatrix} \Phi(z_0, t_1) - z_1 \\ \vdots \\ \Phi(z_{m-1}, t_m - t_{m-1}) - z_m \\ \Gamma_{\sigma_1}(z_{\sigma_1}) \\ \vdots \\ \Gamma_{\sigma_{n_e}}(z_{\sigma_{n_e}}) \end{pmatrix}. \tag{2.8}$$

Non-event time points and z_0 are fixed in (2.8). Solutions to $F = 0$ are obtained by using ADMC++'s Taylor series integrator for hybrid vector fields. The different states of the system correspond to the quiescent and active phases, and event detection is state-dependent.

2.2.4 Periodic orbits without events

Root solvers for locating periodic orbits without events is the subject of [50], with implementation in ADMC++ given in [96] and incorporation into optimization algorithms the subject of [15]. In this situation all time and phase space points are free to vary, with the convention that we are working on closed discrete curves ($z_m = z_0$) with initial time point equal to 0 ($t_0 = 0$). Hence $\mathbf{\Delta} = (t_1, \ldots, t_m, z_0, \ldots, z_{m-1})$ and the multiple shooting map $F : \mathbb{R}^{m\cdot(zdim+1)} \to \mathbb{R}^{m\cdot zdim}$ is given by

$$F(\mathbf{t}, \mathbf{z}) = \begin{pmatrix} \Phi(z_0, t_1) - z_1 \\ \Phi(z_1, t_2 - t_1) - z_2 \\ \vdots \\ \Phi(z_{m-1}, t_m - t_{m-1}) - z_0 \end{pmatrix}. \tag{2.9}$$

Here $F = 0$ is an underdetermined system, with any discrete curve solution belonging to a family of solutions obtained by sliding the mesh points along the periodic orbit. The nullspace of $D_{\mathbf{\Delta}}F$ thus contains the linearization of sliding each mesh point along the orbit. Guckenheimer and Meloon [50] show that appending these nullvectors to $D_{\mathbf{\Delta}}F$ results in an invertible matrix for elementary periodic orbits (orbits with a single multiplier of modulus 1). This invertible matrix can then be used in Newton's method for finding roots of F. Further discussion is given in section 2.3. We can also find solutions by using a Poincare section and simple shooting. For some of the bursting neural models we consider, finding solutions to the multiple shooting map by Newton's method appears to be quite sensitive to the choice of mesh points. Using simple shooting in these situations has been successful.

2.2.5 Periodic orbits with events

The definition of F for periodic orbits with events is the multiple shooting map (2.9) together with the event defining equations:

$$F(\mathbf{t},\mathbf{z}) = \begin{pmatrix} \Phi(z_0, t_1) - z_1 \\ \Phi(z_1, t_2 - t_1) - z_2 \\ \vdots \\ \Phi(z_{m-1}, t_m - t_{m-1}) - z_0 \\ \Gamma_{\sigma_1}(z_{e_1}) \\ \vdots \\ \Gamma_{\sigma_{ne}}(z_{n_e}) \end{pmatrix}. \tag{2.10}$$

The event mesh points are already constrained to satisfy the event defining equations Γ_{σ_i}. Thus to make $D_{\mathbf{\Delta}}F$ square, we need only append equations preventing the sliding of non-event points along the periodic orbit. This is done in the same manner as discussed in section 2.2.4.

As for the non-event case, one option for solving $F = 0$ is to use Newton's method. Another way is to first locate the periodic orbit via a Poincare cross-section and simple shooting, and then to locate event points on this orbit by integrating with ADMC++. Implementation details for this latter method are given in section 2.4.

2.3 Objective functions and their gradients

There are various objective functions G we can choose which reflect the goodness of fit of a computed trajectory to data. Gradient-based optimization algorithms require the gradient of G with respect to parameters. This is obtained from $D_\lambda G =$

$D_{\boldsymbol{\Delta}}G \cdot D_{\lambda}\boldsymbol{\Delta}$. $D_{\boldsymbol{\Delta}}G$ is known for a given choice of G, and $D_{\lambda}\boldsymbol{\Delta}$ can be calculated from (2.4) provided that $D_{\boldsymbol{\Delta}}F$ is invertible. Sections 2.3.1 - 2.3.4 give conditions for when this is the case.

A common form for G is

$$G = \frac{1}{2}\|R(\boldsymbol{\Delta})\|_2^2, \tag{2.11}$$

where R is some residual vector. One choice for R is the voltage residuals:

$$R = (v_1 - v_1^{ref}, \ldots, v_m - v_m^{ref})^T. \tag{2.12}$$

In this case (2.11) becomes the sum of the squared voltage residuals. This is what would be used classically. Another choice is to use the discrepancies in event timings:

$$R = (t_{ev_1} - t_{ev_1}^{ref}, \ldots, t_{n_e} - t_{n_e}^{ref})^T. \tag{2.13}$$

The duty cycle of a periodic burster is defined as the fraction of the period for which the cell is in the active state. This is a biologically important characteristic (cf. section 1.2). We can attempt to match the model duty cycle to data by using the difference in duty cycles in the objective function:

$$R = \frac{t_{term} - t_{init}}{T} - \frac{t_{term}^{ref} - t_{init}^{ref}}{T^{ref}}. \tag{2.14}$$

We can also attempt to match just the period or the burst duration:

$$R = T - T^{ref}, \tag{2.15}$$

$$R = (t_{term} - t_{init}) - (t^{ref}_{term} - t^{ref}_{init}). \tag{2.16}$$

2.3.1 IVPs without events

From (2.7) we get the following expression for $D_{\mathbf{\Delta}}F$ for IVPs without events:

$$D_{\mathbf{\Delta}}F = \begin{pmatrix} -Id & & & \\ D_z\Phi(z_1) & -Id & & \\ & \ddots & \ddots & \\ & & D_z\Phi(z_{m-1}) & -Id \end{pmatrix}. \tag{2.17}$$

This matrix is square and always has full rank.

Proposition 1. *Let F be given by (2.7), with $D_{\mathbf{\Delta}}F$ given by (2.17). Then $D_{\mathbf{\Delta}}F$ has full rank.*

Proof. We have a lower triangular matrix with -1 along the diagonal. Hence the matrix is part of the special linear group. □

2.3.2 IVPs with events

Consider IVPs with events. Let σ_i denote event indices for $i = 1, \ldots, n_e$. Then $(t_{\sigma_i}, z_{\sigma_i})$ denote event points and Γ_{σ_i} the defining function for event point z_{σ_i}.

Let $D_j = D_z\Phi(z_{j-1}, t_j - t_{j-1})$ and $\hat{f}_j = f(\Phi(z_{j-1}, t_j - t_{j-1}))$. We then have the following expression for $D_{\mathbf{\Delta}}F$:

$$D_{\mathbf{\Delta}}F = \begin{pmatrix} -Id & & & & & & & & & \\ & \ddots & & & & & & & & \\ & D_{\sigma_1} & -Id & & & & & \hat{f}_{\sigma_1} & & \\ & & \ddots & \ddots & & & & & \ddots & \\ & & & D_{\sigma_{n_e}} & -Id & & & & & \hat{f}_{\sigma_{n_e}} \\ & & & & \ddots & \ddots & & & & \\ & & & & & D_m & -Id & & & \\ & & \nabla\Gamma_{\sigma_1}^T & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & \nabla\Gamma_{\sigma_{n_e}}^T & & & & & \end{pmatrix}. \tag{2.18}$$

Note that $\hat{f}_j = f(z_j)$ for solutions to $F = 0$. (2.18) is a square matrix. The following proposition shows that $D_{\mathbf{\Delta}}F$ has full rank provided events are transverse.

Proposition 2. *Let F be given by (2.8) and $D_{\mathbf{\Delta}}F$ given by (2.18). Consider $D_{\mathbf{\Delta}}F$ at a solution to $F = 0$ where $<\nabla\Gamma_{\sigma_i}, f(z_{\sigma_i})> \neq 0$ for $i = 1, \ldots, n_e$. Then $D_{\mathbf{\Delta}}F$ has full rank.*

Proof. Reorder the rows of (2.18) so that the row containing $\nabla\Gamma_{\sigma_i}^T$ immediately follows the *zdim* rows containing D_{σ_i}. If the first point is an event point, put the $\nabla\Gamma_1^T$ row immediately following the first *zdim* rows. The resulting matrix contains blocks of the form

$$\begin{pmatrix} D_j & -Id & \ldots & f(z_j) \\ & \nabla\Gamma_j^T & & \end{pmatrix}, \tag{2.19}$$

for event points, and blocks of the form

$$\begin{pmatrix} \dots & D_j & -Id & \dots \end{pmatrix} \tag{2.20}$$

for non-event points. Let C_j denote these blocks. For $j = 2, \dots, m$, C_j is given by (2.19) if z_j is an event point and (2.20) if it is not. Because the initial conditions are treated as fixed the entries for the first mesh point are slightly different. Consequently, let

$$C_1 = \begin{pmatrix} -Id & \dots & f(z_1) \\ \nabla\Gamma_1^T & & \end{pmatrix} \tag{2.21}$$

if z_1 is an event point. Otherwise we set

$$C_1 = \begin{pmatrix} -Id & \dots & \end{pmatrix}. \tag{2.22}$$

Note that C_j has the identity element in columns $(j-1) \cdot zdim + 1$ through $j \cdot zdim$. The transversality conditions give that $\nabla\Gamma_j \neq 0$. Thus each row of C_j has a nonzero entry in the columns $(j-1) \cdot zdim + 1$ to $j \cdot zdim$. Note also that the entries of C_j in columns $j \cdot zdim + 1$ through $m \cdot zdim$ are all zeros.

Next we show that the rows of C_j are independent for all j. This is immediate if z_j is not an event point due to the identity block. So let z_j be an event point. The first $zdim$ rows of (2.19) are independent because of the identity block. We show that the last row is independent of the first $zdim$ rows by contradiction. Let $r_1, \dots, r_{zdim+1}$ denote the rows of (2.19). Suppose $\exists \alpha_\ell$ not all zero such that $\sum_{\ell=1}^{zdim} \alpha_\ell r_\ell + r_{zdim+1} = 0$. Then because of the identity block, we have $\alpha_\ell = (\nabla\Gamma_j)_\ell$. As r_{zdim+1} has all zeros in the t_j column, we have $\sum_{\ell=1}^{zdim} \alpha_\ell f_\ell = 0 \Rightarrow \sum_{\ell=1}^{zdim} (\nabla\Gamma_j)_\ell \cdot f_\ell = 0$. In other words, $<\nabla\Gamma_j, f(z_j)> = 0$. But this has been ruled out by the transversality assumption.

We now show that (2.18) has full rank. We have seen that the rows of C_1 are independent. Assume that the rows of $C_1, \ldots, C_k$ are independent for some $k \in \{1, \ldots, m-1\}$. We show that this implies that the rows of $C_1, \ldots, C_{k+1}$ are independent. The rows of C_{k+1} are independent of one another, so it suffices to show that each row of C_{k+1} is independent of the rows of $C_1, \ldots, C_k$. But this is immediate as each row of C_{k+1} has a nonzero entry in the columns $k \cdot zdim + 1$ through $(k+1) \cdot zdim$, whereas all entries of $C_1, \ldots, C_k$ in these columns are zero. □

2.3.3 Periodic orbits without events

F is given by (2.9) for periodic orbits without events. Let $D_i = D_z\Phi(z_i, t_{i+1} - t_i)$, and let $\hat{f}_i = f(\Phi(z_{i-1}, t_i - t_{i-1}))$. Then $D_{\mathbf{\Delta}}F$ is given by

$$D_{\mathbf{\Delta}}F = \begin{pmatrix} D_0 & -Id & & & \hat{f}_1 & & \\ & \ddots & \ddots & & & \ddots & \\ & & D_{m-2} & -Id & & & \hat{f}_{m-1} \\ -Id & & & D_{m-1} & & & \hat{f}_m \end{pmatrix}. \tag{2.23}$$

When evaluated at solutions to $F = 0$, $\hat{f}_i = f(z_i)$ the vector field evaluated at z_i. The kernel of $D_{\mathbf{\Delta}}F$ contains the vectors corresponding to the linearization of sliding mesh points along the periodic orbit:

$$\begin{aligned} v_0 &= \big(f(z_0)^T, 0, \ldots, -1, -1, \ldots, -1, 0\big)^T, \\ v_1 &= \big(0, f(z_1)^T, \ldots, 0, 1, 0, \ldots, 0\big)^T, \\ &\vdots \\ v_{m-1} &= \big(0, \ldots, f(z_{m-1})^T, 0, \ldots, 1, 0\big)^T. \end{aligned} \tag{2.24}$$

Appending these vectors as rows to (2.23) gives a square matrix. Guckenheimer and Meloon [50] prove that $D_{\mathbf{\Delta}}F$ has full rank for periodic orbits possessing a single multiplier equal to one.

Proposition 3. *Consider a periodic orbit without events which has a single multiplier equal to 1. Let $D_{\mathbf{\Delta}}F$ be given by (2.23) evaluated at the periodic orbit, together with the appended vectors (2.24). Then $D_{\mathbf{\Delta}}F$ has full rank.*

Proof. See [50]. □

For elementary periodic orbits we can thus invert $D_{\mathbf{\Delta}}F$ both to obtain $D_\lambda \mathbf{\Delta}$ from (2.4), and for use in Newton's method for locating solutions to $F = 0$. Newton steps $\mathbf{\Delta}_s$ satisfy $F(\mathbf{\Delta}_c) = D_{\mathbf{\Delta}}F(\mathbf{\Delta}_c) \cdot \mathbf{\Delta}_s$. The (augmented) last m components of $F(\mathbf{\Delta}_c)$ are set equal to 0, so the Newton steps are orthogonal to the motion along the periodic orbit at the mesh points.

2.3.4 Periodic orbits with events

Consider $D_{\mathbf{\Delta}}F$ for periodic orbits with events. Starting again with (2.23), we need phase conditions for the mesh points to get a square matrix $D_{\mathbf{\Delta}}F$. Phase conditions for the event points are given by the event defining equations. The remaining phase conditions are given by appending the tangent vectors which correspond to sliding these non-event mesh points along the orbit, as discussed in section 2.3.3. Letting $D_i = \Phi(z_i, t_{i+1} - t_i)$ and $\hat{f}_i = f(\Phi(z_{i-1}, t_i - t_{i-1}))$ we have

$$D_{\mathbf{\Delta}}F = \begin{pmatrix} D_0 & -Id & & & & \hat{f}_1 & \\ & \ddots & & & & \ddots & \\ -Id & & & & D_{m-1} & & \hat{f}_m \\ & \nabla\Gamma_{\sigma_1}^T & & & & & \\ & & \ddots & & & & \\ & & & \nabla\Gamma_{\sigma_{n_e}}^T & & & \\ v_{\rho_1}^T & & & & & & \\ & & \ddots & & & & \\ & & & & v_{\rho_{m-n_e}}^T & & \end{pmatrix}. \tag{2.25}$$

σ_i denotes event point indices for $i = 1, \ldots, n_e$, and non-event point indices are given by ρ_j, $j = 1, \ldots, m - n_e$. The vectors v_{ρ_j} are the elements of (2.24) corresponding to non-event points. The following proposition shows that $D_{\mathbf{\Delta}}F$ has full rank if evaluated at an elementary periodic orbit with transverse events.

Proposition 4. *Let F be given by (2.10), and $D_{\mathbf{\Delta}}F$ by (2.25). If $D_{\mathbf{\Delta}}F$ is evaluated at a periodic orbit with a single multiplier of modulus 1, and $<\nabla\Gamma_{\sigma_i}, f(z_{\sigma_i})> \neq 0$ for $i = 1, \ldots, n_e$, then $D_{\mathbf{\Delta}}F$ has full rank.*

Proof. Let A denote the submatrix given by the first $m \cdot zdim$ rows of (2.25), B consist of rows $m \cdot zdim + 1$ through $m \cdot zdim + n_e$, and C denote the last $m - n_e$ rows. B corresponds to the rows of $D_{\mathbf{\Delta}}F$ containing the $\nabla\Gamma_{\sigma_i}$ terms, and C the v_{ρ_j} elements of (2.25).

Guckenheimer and Meloon [50] show that A has full rank for elementary periodic orbits. Because $<\nabla\Gamma_{\sigma_i}, f(z_{\sigma_i})> \neq 0$, $\nabla\Gamma_{\sigma_i} \neq 0$ and the rows of B are therefore

linearly independent. We next show that transverse events imply that each row of B has a nonzero component in $N(A)$. The inner product of $\nabla\Gamma_{\sigma_i}$ with v_{σ_i}, the basis element of $N(A)$ corresponding to sliding mesh point σ_i along the orbit, is equal to $<f(z_{\sigma_j}), \nabla\Gamma_{\sigma_j}>$. This is nonzero by the transversality assumption. Let $[A; B]$ denote the first $m \cdot zdim + n_e$ rows of $D_{\mathbf{\Delta}}F$. We thus have that $[A; B]$ has full rank. Finally, note that a basis for the nullspace of $[A; B]$ is given by the basis elements of $N(A)$ corresponding to sliding non-event points along the periodic orbit. These basis elements are the rows of C. □

2.4 Taking the optimization step

Our objective functions G are functions of solution trajectories $\mathbf{\Delta}$ of a particular type, for example periodic bursting orbits. The $\mathbf{\Delta}$ at which we evaluate G is defined as a solution to $F(\mathbf{\Delta}) = 0$. Solutions $\mathbf{\Delta}(\lambda)$ depend upon parameters $\lambda \in \mathbb{R}^p$. Our objective function is undefined if we leave the domain $\mathcal{D} \subset \mathbb{R}^p$ where solutions to $F = 0$ exist. For IVPs without events this is not a problem, but for the other problem types this is a concern. Boundaries for the existence of periodic orbits or bursting are not typically known beforehand. Location of periodic orbit boundaries for many active parameters is equivalent to computing multi-dimensional bifurcation surfaces, which is a challenging task. Thus we need methods to determine whether a proposed parameter step leaves the feasible domain, as well as ways to prevent these steps from being taken. We are confining ourselves to interior point methods here due to the fact that G is undefined outside of the valid domain.

We use an infinite barrier function and set $G = \infty$ outside of the feasible domain $\mathcal{D}$. This is a crude method and there is obviously much room for refinement here. To help illustrate how our algorithm works for determining whether a parameter

step leaves the domain, we walk through the steps of Algorithm 1 for a bursting periodic orbit problem.

Algorithm 1 General optimization algorithm

Initialize current parameters λ_c, parameter step λ_s

while termination criteria not met **do**

Set $\lambda_+ \leftarrow \lambda_c + \lambda_s$

Solve $F(\mathbf{\Delta}(\lambda_+), \lambda_+) = 0$

Compute $D_\lambda \mathbf{\Delta}$

Compute $G(\mathbf{\Delta}(\lambda_+)), D_\lambda G$

if λ_s is acceptable **then**

$\lambda_c \leftarrow \lambda_+$

end if

Compute new parameter step λ_s

end while

2.4.1 Steps for a bursting periodic orbit

Suppose we have periodic bursting data that we would like to fit using some gradient-based method. Let λ_c be the current parameter values and λ_s a proposed parameter step. The process of finding an acceptable parameter step and calculating a proposed subsequent step is described below.

1. **Solve $F = 0$.** First we attempt to find $\mathbf{\Delta}$ satisfying $F(\mathbf{\Delta}) = 0$ for the proposed parameters $\lambda_+ = \lambda_c + \lambda_s$. This is accomplished by the following algorithm.

1.a **Generate an initial guess $\mathbf{\Delta}_0$.** If the parameter step is large the previously computed solution trajectory $\mathbf{\Delta}$ may not be a good initial guess. We are working with stable periodic orbits here, so we integrate forwards in time to get close to the stable orbit.

1.a.i **Find silent phase.** The silent phase is located by taking a starting point, assuming that it is in the active phase, and integrating forwards in time until either a termination event is detected or a maximum detectable burst period is reached. No termination event will be detected if trajectories for λ_+ converge to a fixed point, or if λ_+ corresponds to a periodic burster whose period is larger than our set tolerance. Setting a maximum detectable burst period is difficult because periodic bursters can have unbounded period (for example, near homoclinics and SNICs). If λ_+ appears to lie outside the feasible domain, we set $G = \infty$ and go to step 2.

1.a.ii **Set a Poincare cross-section.** We use simple shooting and a Poincare cross-section to locate the periodic orbit. For the models we have used, a cross-section based upon a slow variable value during the silent phase has worked well. Let y denote a slow variable. We select a point in the silent phase where $|\dot{y}|$ is sufficiently away from zero to help ensure return to the cross-section.

1.b **Locate orbit by simple shooting.** For the examples considered in this thesis, simple shooting using Newton's method and finite difference approximations has worked well. There are alternatives including multiple shooting. Failure of Newton's method to converge suggests a non-periodic orbit, for example a chaotic burster, in which case we set

$G = \infty$.

1.c **Event location and computation of DF.** Given a periodic orbit, we locate events by integrating once around the orbit using ADMC++'s Taylor series integrator for hybrid systems and monitoring for events. ADMC++ also computes the Jacobian of the flow between mesh points, which allows us to compute $D_{\mathbf{\Delta}} F$ and thus $D_\lambda \mathbf{\Delta}$.

2. **Accept / reject parameter step.** Having obtained $\mathbf{\Delta}$ satisfying $F(\mathbf{\Delta}) = 0$, we compute our objective function $G(\mathbf{\Delta}(\lambda))$. Whether the parameter step is accepted or rejected is typically based upon comparing $G(\lambda_+)$ with $G(\lambda_c)$. For example, descent algorithms typically require $G(\lambda_+) < G(\lambda_c)$. Comparisons involving derivatives of G may also be involved. For example, a trust region method may expand the trust radius and attempt a larger step if there is a good match between $G(\lambda_+)$ and the predicted value from the quadratic model of G at λ_c.

3. **Compute new parameter step.** A new parameter step is then calculated. This is based upon whether the previous parameter step was accepted or rejected. For example, we may shrink the parameter step if the previous parameter step was rejected because G did not decrease. New parameter steps may also depend upon derivatives of G. For example, following a successful parameter step the next parameter step may be calculated as the minimizer of a quadratic model of G, where the quadratic model depends upon first and second order derivatives of G.

Chapter 3

Test problems

This chapter considers test problems drawn from two systems. The first is a minimal spiking model studied in Guckenheimer et al [51] as a generic fast subsystem for bursters. We use this model to show the importance of automatic differentiation. The idea of the minimal model is to absorb slow currents into an effective leak current. This approach will be relevant in our study of preBotzinger respiratory neurons (chapter 4). The second system is the Hindmarsh-Rose equations, a polynomial vector field which exhibits an array of behavior including bursting. We use this system as a test bed for our bursting optimization algorithms. In the process we encounter some delicate behavior which poses numerical challenges.

3.1 A minimal spiking model

This section discusses a minimal spiking model that was studied in [51]. We use this model to illustrate that the increased accuracy of automatic differentiation over finite differences can be important to optimization algorithm performance. The model is also of interest in its own right in regards to bursting. The equations represent a fast subsystem which comes from considering fast and slow *currents*, as opposed to fast and slow phase space variables as is traditionally done [104]. Slow currents are absorbed into an effective leak current. Bifurcation analysis of the resulting fast subsystem with the effective leak conductance and reversal potential as the bifurcation parameters gives insight into burst initiation and termination mechanisms.

3.1.1 Motivation and bifurcations of the minimal model

In a Hodgkin-Huxley type model, currents for a particular channel type are represented as

$$I_i = \bar{g}_i m_i^p h_i^q (v - v_i). \tag{3.1}$$

Currents are ohmic, the product of a conductance and an electrical potential. The cell's membrane potential is given by v, and v_i is the channel's reversal potential. The conductance g of a channel is the product of its maximal conductance $\bar{g}$ with the activation and inactivation gating variables m and h. If the kinetics of the gating variables are sufficiently slow, we say that the current is a slow current. For given values of g_i, m_i, and h_i for the slow currents, we can absorb the slow currents into an effective leak current:

$$\begin{aligned} g_{L_{\text{eff}}} &= \sum g_i, \\ v_{L_{\text{eff}}} &= \sum \frac{g_i}{g_{L_{\text{eff}}}} v_i. \end{aligned} \tag{3.2}$$

The effective leak conductance is the sum of the slow conductances. The effective leak reversal potential is the weighted average of the slow reversal potentials, with the weights proportional to the individual channel conductances. Absorbing all currents other than a fast sodium current into an effective leak current gives the minimal spiking model studied by Guckenheimer et al [51]:

$$
\begin{aligned}
\dot{v} &= -(g_{\mathrm{Na}} m^3 h(v - v_{\mathrm{Na}}) + g_{\mathrm{L}}(v - v_{\mathrm{L}})) \\
\dot{h} &= \alpha_h(v)(1-h) - \beta_h(v)h \\
\alpha_h(v) &= 0.07 \exp(\frac{-60 - v}{20}) \\
\beta_h(v) &= \frac{1}{1 + \exp(\frac{-30 - v}{10})} \\
m &= \frac{1}{1 + \exp(\frac{-(v+35)}{9})}
\end{aligned}
\tag{3.3}
$$

System (3.3) can be viewed as a fast subsystem, where $(g_{L_{\mathrm{eff}}}, v_{L_{\mathrm{eff}}})$ evolve on a slower time scale. We can gain insight into burst initiation and termination mechanisms by examining bifurcations which occur in the $(g_{L_{\mathrm{eff}}}, v_{L_{\mathrm{eff}}})$ plane. The most relevant bifurcations of system (3.3) in the context of optimization are shown in Figure 3.1. There exists a tear-shaped region whose boundaries consist of supercritical Hopf points and saddle-nodes of limit cycles. Stable periodic orbits exist within this region. A more complete bifurcation diagram is given in [51], and some interesting codimension-2 bifurcations are discussed.

This strategy of absorbing slow currents into an effective leak current can be applied to a variety of bursting models. Which currents to include in the effective leak current is a question that must be determined for each model. Both $g_{L_{\mathrm{eff}}}$ and $v_{L_{\mathrm{eff}}}$ must change sufficiently slowly in order to have a good separation of time scales. It is useful to see that $\dot{v}_{L_{\mathrm{eff}}}$ depends not only upon $\dot{g}_i$ but also upon how different the constituent reversal potentials are from one another:

$$
\dot{v}_{L_{\mathrm{eff}}} = \sum \frac{\dot{g}_i}{g_{L_{\mathrm{eff}}}} (v_i - v_{L_{\mathrm{eff}}}). \tag{3.4}
$$

Even a small value of $|\dot{g}_i|$ can lead to a large change in $\dot{v}_{L_{\mathrm{eff}}}$ if $|v_i - v_{L_{\mathrm{eff}}}|$ is large. Similar values of v_i will tend to result in small values of $\dot{v}_{L_{\mathrm{eff}}}$. A case in point is

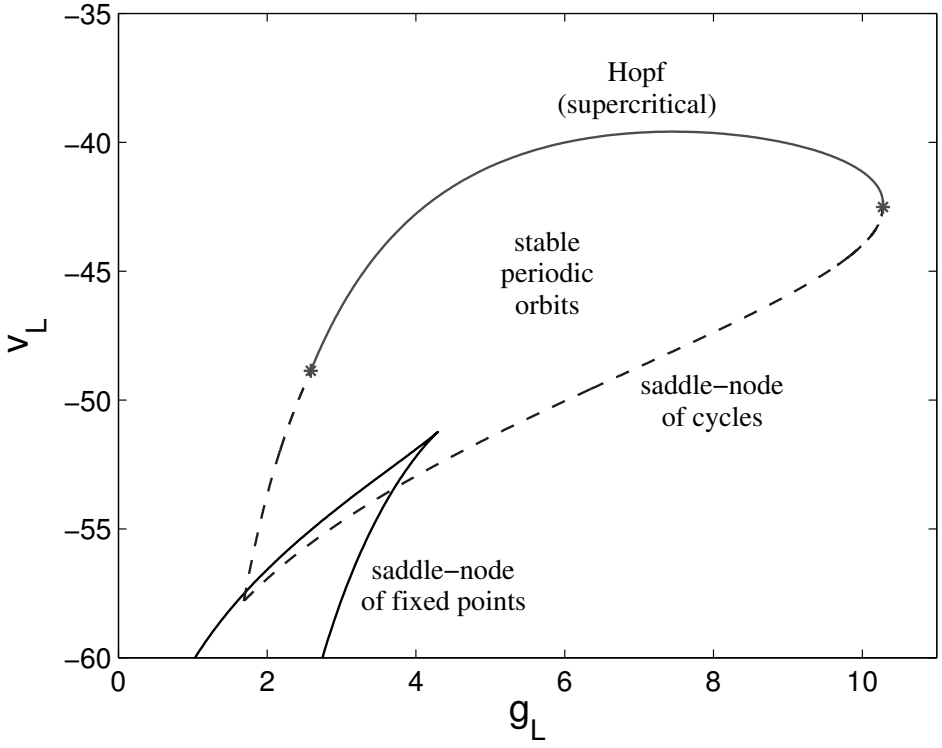


Figure 3.1: A partial bifurcation diagram for system (3.3). A tear-shaped region is formed from a curve of supercritical Hopf points together with saddle-nodes of limit cycles (dashed curve). Stable periodic orbits exist within this region. The saddle-nodes of cycles terminate at generalized Hopf points (asterisks). Also shown are saddle-nodes of fixed points, including a codimension-2 cusp point.

Model IV of [19], a model of leech heart interneurons which displays plateau oscillations. The currents involved are a fast sodium current, a persistent potassium current, and a leak current. The persistent potassium current has slow kinetics and a reversal potential fairly close to the leak's reversal potential. Combining the two into a single effective leak current results in a good separation of time scales. Burst initiation and termination correspond closely to crossing bifurcation curves in the $(g_{L_{\text{eff}}}, v_{L_{\text{eff}}})$ plane. Further discussion of this and the effective leak current reduction in several other models is given in Guckenheimer et al [51].

3.1.2 Optimization and the minimal model

We generate a reference trajectory from (3.3) with $(g_{\text{L}}, v_{\text{L}}) = (9.6493, -59.2293)$. This corresponds to a quiescent cell, so the reference trajectory approaches a stable fixed point. Let our objective function G be given by the sum of the voltage squared residuals. The resulting objective function landscape is shown in Figure 3.2, together with the teardrop within which stable periodic orbits exist.

There are two features of the landscape that are of particular interest. First is that the qualitative features of solution trajectories are clearly reflected in the landscape. The wrinkles within the teardrop reflect that within this region computed trajectories are periodic, while the reference trajectory is quiescent. The appearance of new wrinkles as we move across the teardrop by increasing v_{L} corresponds to the appearance of new spikes in the voltage trace. As we increase v_{L}, more spikes appear in our fixed time span because the period of the orbit is decreasing. The amplitude of the periodic orbits is also decreasing in this direction, explaining why the height of the wrinkles decrease as we increase v_{L}. As we increase v_{L} past the boundary of the teardrop, we cross a Hopf bifurcation where

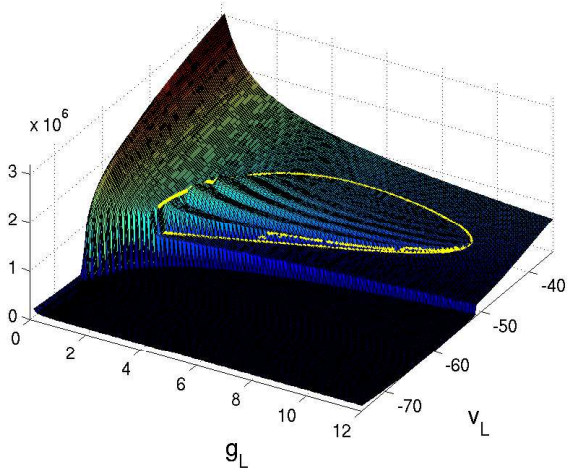


Figure 3.2: Objective function landscape for system (3.3), with reference trajectory corresponding to $(g_{\mathrm{L}}, v_{\mathrm{L}}) = (9.6493, -59.2293)$ and objective function given by $G = \sum_i (v_i - v_i^{ref})^2$. The loop denotes the tear-shaped region where stable periodic orbits exist (see Figure 3.1).

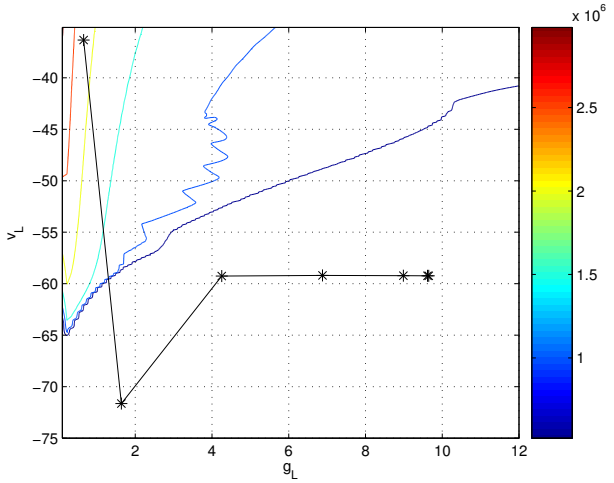


Figure 3.3: A contour map of the landscape in Figure 3.2, together with parameter steps taken by using Levenberg-Marquardt together with automatic differentiation and Taylor series integration. The algorithm converges to the global minimum at $(g_{\mathrm{L}}, v_{\mathrm{L}}) = (9.6493, -59.2293)$. Automatic differentiation is necessary for convergence because the region surrounding the global minimum is nearly flat.

the amplitude of the periodic orbits shrinks to zero and the wrinkles disappear.

The wrinkles are liable to lead to local minima. For example, this will be the case if we optimize in v_{L} for a fixed g_{L} which transects the teardrop. This suggests that optimization algorithm performance is likely to be improved if we take into account the qualitative features we are trying to capture, for example whether the reference data is periodic. Restricting to the appropriate class of problem (e.g. periodic orbits) should lead to nicer landscapes. Casey [15] developed these ideas for periodic orbits without events.

The other prominent feature of the landscape in Figure 3.2 is the nearly flat "plain" for more negative v_{L} values. This region corresponds to quiescent cells. Evidently, quiescent trajectories for system (3.3) are insensitive to the leak parameters. Within the plain, the partial derivatives of trajectory points with respect to the leak parameters have magnitude $\approx 2 \times 10^{-7}$. This is at or beyond the limit of accuracy for computing derivatives using finite differences. Using the Levenberg-Marquardt algorithm with finite differences leads to premature termination away from the global minimum. On the other hand, this level of accuracy is well within the reach of automatic differentiation. Using Levenberg-Marquardt with automatic differentiation leads to convergence to the global minimum. The parameter steps taken are shown in Figure 3.3. Hayes et al [54] report similar flat landscapes in quiescent parameter regions.

3.2 A bursting example: Hindmarsh-Rose

This section uses the Hindmarsh-Rose equations [57, 58] to illustrate some of the ideas developed in chapter 2. In particular, we focus on burst trajectories. Both transient (non-periodic) and periodic bursting are considered. The multiple time

scales involved in bursting lead to complex phenomena that make fitting burst trajectories a challenging problem. This section illustrates some of these challenges, and shows how using burst events as defined in (2.5)-(2.6) can improve algorithm performance.

3.2.1 Hindmarsh-Rose equations

The Hindmarsh-Rose equations (3.5) are a well-known system of ODEs which were introduced to describe various firing patterns in pond snail neurons [57, 58]. The motivation for this polynomial vector field was to revise Fitzhugh's model [35] in order to capture certain observed features such as tail currents [56]. Like Fitzhugh's model, the Hindmarsh-Rose equations are mathematically more tractable than Hodgkin-Huxley style models, yet still display a rich array of behaviors, including tonic spiking, bursting, and chaotic firing patterns. As a result, the Hindmarsh-Rose equations have been studied in a variety of contexts ranging from neural systems (e.g. [12, 60, 100, 112]) to chemical oscillations [20, 93] to control of chaotic trajectories [106]. We use the Hindmarsh-Rose equations for testing our algorithms because it is a computationally inexpensive system to solve which displays all of the qualitative features we are interested in here.

$$\begin{aligned} \dot{x} &= y - ax^3 + bx^2 + I - z \\ \dot{y} &= c - dx^2 - y \\ \dot{z} &= r\big(S(x - x_1) - z\big) \end{aligned} \tag{3.5}$$

The phase space variable x represents the cell's membrane potential, y is a recovery variable, and z denotes an injected current which varies with time. The parameter I represents a constant external injected current, while r governs the

z time scale. Typically $r \ll 1$, in which case x, y are the fast variables and z the slow variable. In this case $I_{eff} = I - z$ denotes a slowly varying effective injected current. We will use I_{eff} as a bifurcation parameter when analyzing the fast subsystem dynamics. Default parameter values are $a = 1, b = 3, c = 1, d = 5, x_1 = -1.6, S = 4$. I and r will be varied throughout this section.

We focus on burst trajectories of (3.5). One of the main features of a burst trajectory is the burst duration. This in turn is affected by the number of spikes in the burst. For periodic bursts, the period and burst duration are also linked with one another. It turns out that the number of spikes in a burst depends very sensitively on the parameters [117]. We illustrate this in the next section.

3.2.2 Continuation and canards

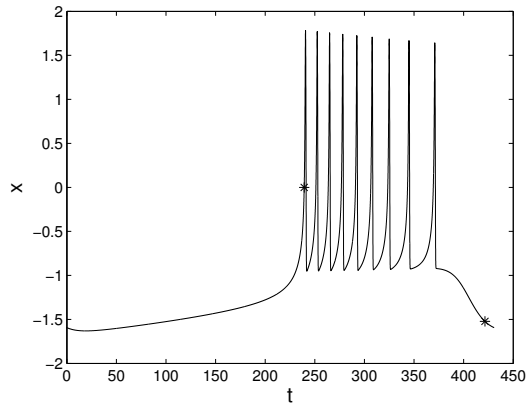
For $I = 2, r = 1 \times 10^{-3}$, a periodic burst solution exists for (3.5). This trajectory is shown in Figure 3.4. Detected event points using (2.5)-(2.6) with $\delta = 1 \times 10^{-2}$ are marked by asterisks.

We use a simple shooting method to continue this family of periodic orbits while keeping the period fixed. This corresponds to computing zero sets of the defining equation

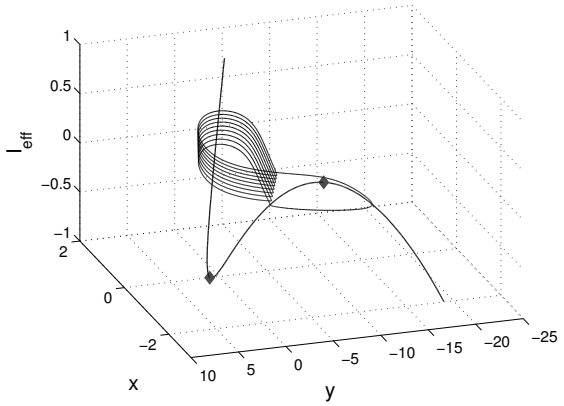
$$F_T(u_0, \lambda) = \Phi(u_0, T, \lambda) - u_0. \tag{3.6}$$

Here T denotes the fixed period of the orbits.

The Jacobian of (3.6) does not have full rank, as the periodic orbit's tangent directions lie in the nullspace. To obtain a full rank map, we take our initial point u_0 on the periodic orbit and append the vector $(f(u_0), 0, \ldots, 0)$ to the Jacobian in the same fashion as for our multiple shooting maps (section 2.2). Provided we



(a)



(b)

Figure 3.4: Burst trajectory for the Hindmarsh-Rose system (3.5) with $I = 2, r = 1 \times 10^{-3}$. (a) Event points located with local event checking and $\delta = 1 \times 10^{-2}$ are marked by asterisks. (b) Solution curve together with the critical manifold M_0. Fold points of M_0 are marked by diamonds. The component of M_0 between the fold points consist of saddle-points of the fast subsystem. Note that burst termination is associated with a close passage to the saddle component of M_0. The effective value of I is given by $I_{eff} = I - z$.

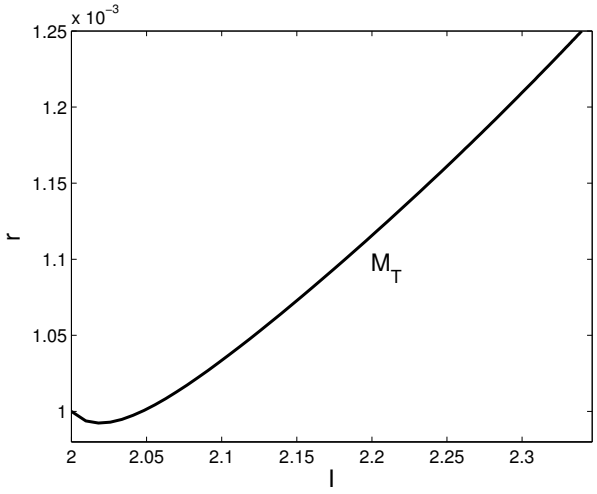
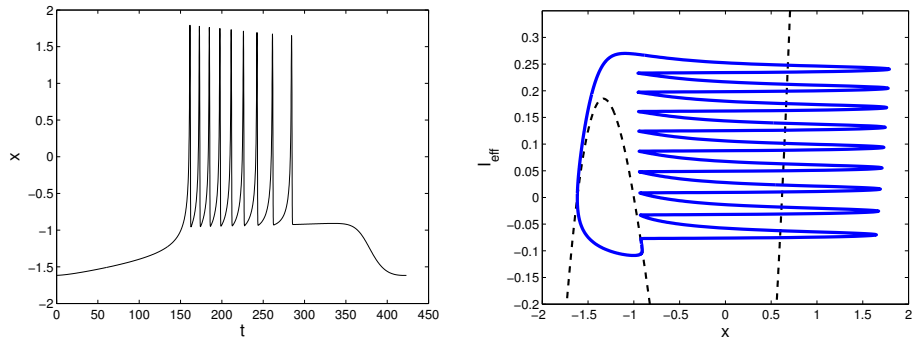


Figure 3.5: Curve of fixed period bursters $\mathcal{M}_T$ in the (I, r) plane. Period along the curve is 430.78.

have an elementary periodic orbit [50], appending an additional component to F_T which is identically zero gives a full rank map from $\mathbb{R}^{n+ap} \to \mathbb{R}^{n+1}$, where ap is the number of active parameters. Thus continuing a curve of fixed period orbits requires two active parameters.

We use a predictor-corrector continuation algorithm with natural corrector steps, where corrector steps are orthogonal to the most rapidly varying parameter in the predictor step [75]. Taking I and r as the active continuation parameters and starting from $(I, r) = (2, 1 \times 10^{-3})$, we obtain a curve of fixed period bursters in the (I, r) plane (Figure 3.5). Along this curve, the burst duration varies but the spike number does not change. Figure 3.6 shows a member of this family where the burst duration has increased significantly. This burster shows a voltage plateau following the last spike in the burst. The voltage "hangs" close to threshold for a significant length of slow time before returning to the quiescent state.

The voltage plateau corresponds to a *canard*, where the solution trajectory follows an unstable manifold for $\mathcal{O}(1)$ of slow time [25]. This is illustrated in



(a) (b)

Figure 3.6: Periodic burst trajectory for the Hindmarsh-Rose system (3.5) which shows canards. $I = 2.28409796167839, r = 1.19380932644 \times 10^{-3}$. (a) Plot of x vs. t. (b) Periodic burst orbit (solid) together with the critical manifold M_0 (dashed) projected to the $x - I_{eff}$ plane. The plateau in (a) corresponds to the portion of the orbit which follows the middle (saddle) branch of M_0.

Figure 3.6b, which shows the periodic orbit together with the critical manifold M_0 (the set of fixed points of the fast subsystem). The segment of M_0 between the two fold points consists of fast subsystem saddle points. The long plateau at the end of the active phase of the burst corresponds to following this unstable saddle segment of M_0. Following an unstable manifold is a very delicate phenomenon that is easily destroyed by slight changes in parameter values. Figure 3.7(a)-(b) shows the periodic orbit after I has been perturbed by a tenth of a percent. The resulting burst trajectory has 10 spikes rather than the original 9. Similarly, if we do not change the parameter values but change the numerical integration tolerances to 1×10^{-5} from the original 1×10^{-8}, we get a trajectory which once again fires 10 spikes rather than 9. Furthermore, the new trajectory is clearly non-periodic. See Figure 3.7(c)-(d).

Canards were originally discovered and analyzed in the van der Pol equation [25, 26, 49, 123]. The canard orbits occur in exponentially thin regions of parameter

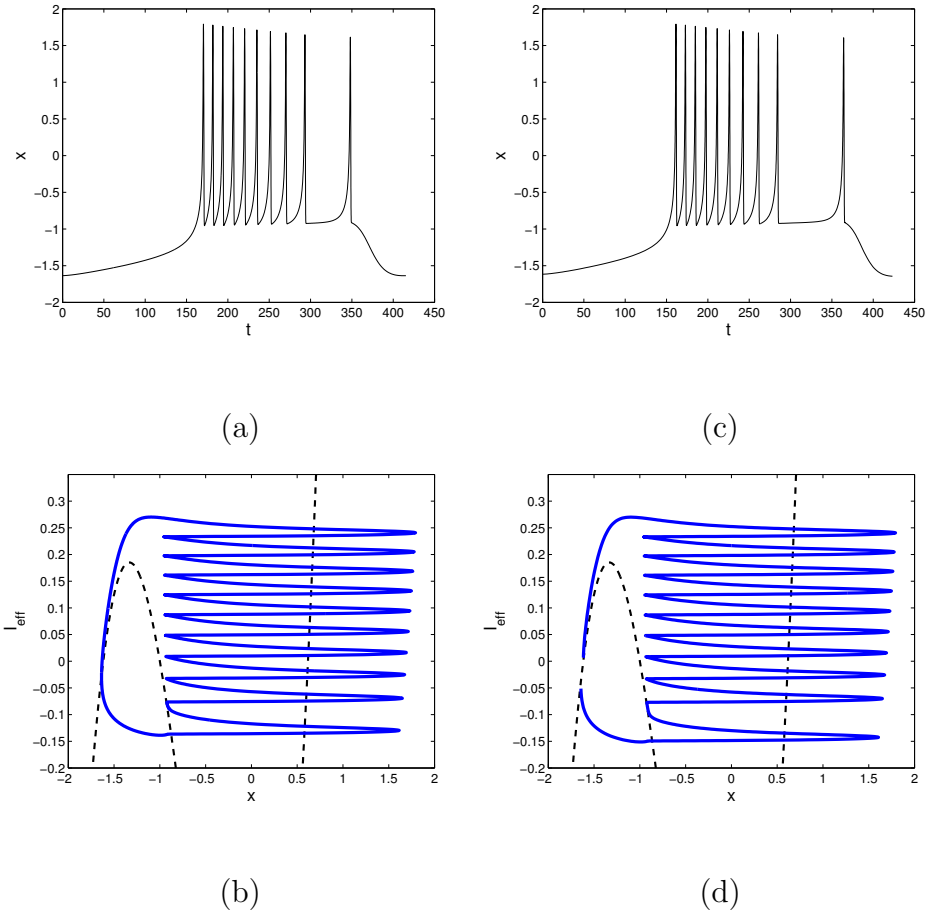


Figure 3.7: Perturbations of the canard orbit in Figure 3.6 result in an additional spike. (a)-(b) I=2.28638205964007. All other parameters the same as in Figure 3.6. (c)-(d) Same parameters as in Figure 3.6, but Matlab's `ode15s` solver is used with absolute and relative tolerances equal to 1×10^{-5}. From (d) we see that the computed trajectory is not periodic.

space [26]. Based on this, the delicacy of the burst orbit shown in Figure 3.6 is not surprising. Slight changes in the parameters result in different spike numbers. Terman [117] examined transitions in spike number for a certain class of burst periodic orbits. The transitions studied by Terman [117] either occurred through chaotic transitions, or occurred smoothly via canard trajectories. We have not pinned down the exact mechanism by which spike number changes in system (3.5). What we can say is that these changes occur in thin regions of parameter space, and that nearby trajectories involve canards.

3.2.3 Optimization algorithm performance: fitting a transient burst

When the number of spikes in a burst trajectory changes, the burst duration changes as well. This is evident in Figures 3.6. Plateau trajectories of increasing length occur leading up to the addition of a spike. The new spike is added at the end of this plateau. As the plateau corresponds to a canard, it is very sensitive to parameter changes. Thus the final interspike interval shrinks rapidly following addition of the new spike, and spike addition results in first an increase and then a decrease in burst duration. For periodic bursters, there is a corresponding increase in period leading up to spike addition, and then a decrease in period immediately afterwards. This can lead to multiple local minima in objective function landscapes.

Discrepancies in finer details such as spike location may still exist even when gross features such as burst duration and period match well. Changing spike timing within a fixed burst window can lead to multiple local minima. Objective function landscapes are thus expected to be complicated. We illustrate this by trying to fit

a transient burster from system (3.5).

For the remainder of this section, we use the voltage time trace for $t \in [0, 500]$ from system (3.5) with $(I_{ref}, r_{ref}) = (0.4, 1\times10^{-3})$ and initial conditions $(-1.6, -10, 0)$ as a reference trajectory. The reference trajectory is a transient burst with 8 spikes and is shown in Figure 2.2. Because there are several spikes in the reference burst, there is ample opportunity for local minima due to shifting alignment of the spikes. This is reflected in the many wrinkles and grooves in Figure 3.8, which plots the objective function landscape using the sum of voltage squared residuals. The complications in the landscape also reflect changes in spike number. This is illustrated in Figures 3.10(a)-(b), where a transect of the landscape for fixed I is plotted against r, together with the number of spikes plotted against r. We see that for small values of r, local minima mostly correspond to changes in spike number, whereas for larger values of r the local minima reflect both changes in spike number and spike timing.

The burst event functions (2.5)-(2.6) capture coarse bursting features. Ignoring the trajectory details at non-event points and attempting to match burst event timing alone should avoid some of the local minima associated with spike timing within the burst. Figure 3.9 shows the resulting landscape for $G = \frac{1}{2}\big((t_{init} - t_{init}^{ref})^2 + (t_{term} - t_{term}^{ref})^2\big)$ using local event checking. The landscape is much less complicated than in Figure 3.8. There are still local minima due to changes in spike number. Whether spike addition is associated with a local minimum or maximum depends upon whether t_{term} is bigger or smaller than t_{term}^{ref}. To see this, consider the transect shown in Figure 3.10c. Spike number decreases with r. Consider first the case where $t_{term} > t_{term}^{ref}$. This tends to be the situation when the number of spikes is larger than the number in the reference trajectory. As we increase

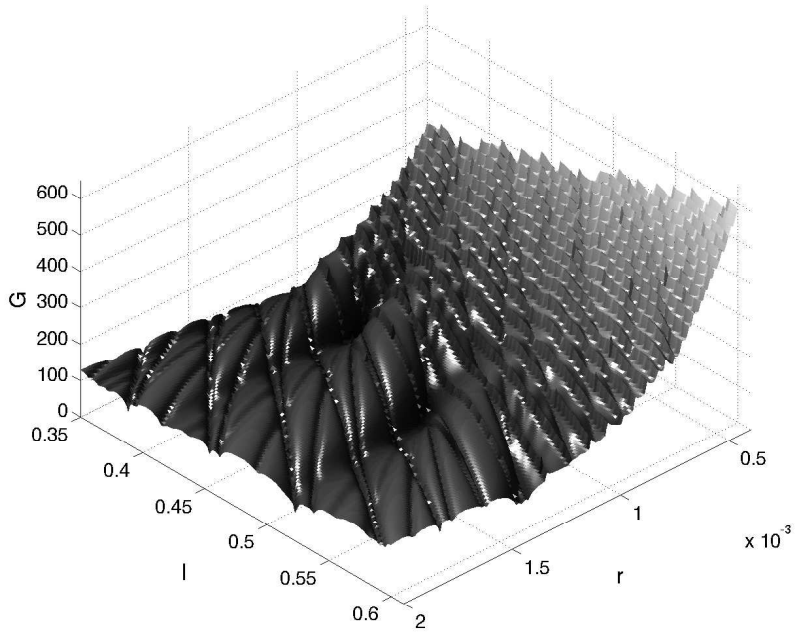


Figure 3.8: Objective function landscape for system (3.5) with a reference trajectory $(I_{ref}, r_{ref}) = (0.4, 1 \times 10^{-3})$ a transient burst. 10,000 time mesh points were evenly spaced on $[0, 500]$. The objective function was given by the inter-mesh interval $\times$ the sum of the voltage (x) squared residuals. This can also be viewed as an approximation to the L_2 distance between the voltage curves. $a = 1, b = 3, c = 1, d = 5, x_1 = -1.6, S = 4$.

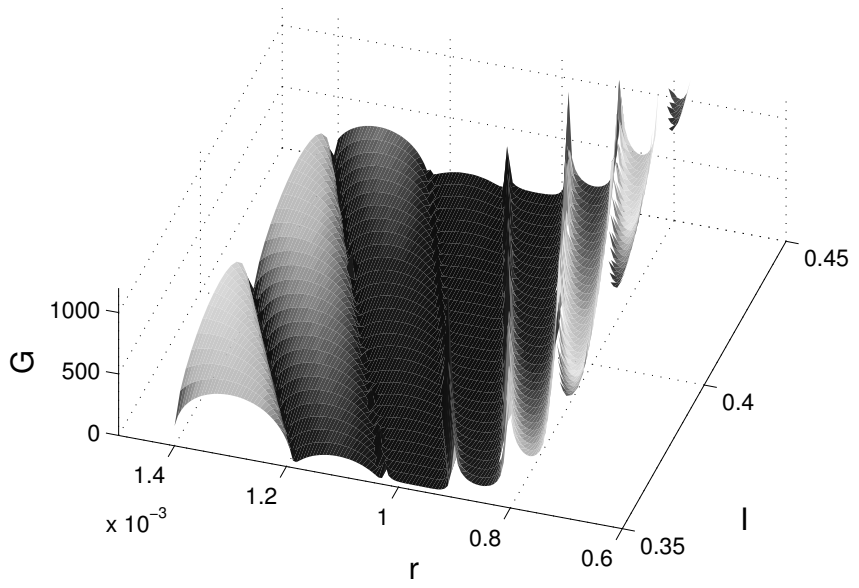


Figure 3.9: Event timing objective function landscape for system (3.5) with a reference trajectory $(I_{ref}, r_{ref}) = (0.4, 1 \times 10^{-3})$ a transient burst. Local event checking was used (cf. section 2.2.1). $a = 1, b = 3, c = 1, d = 5, x_1 = -1.6, S = 4$.

r towards the point where the spike number changes, the last interspike interval increases, thus increasing t_{term}. Termination event timing is much more sensitive to r than initiation timing is ($|\frac{\partial t_{init}}{\partial r} \cdot \bar{r}| = 2.26 \times 10^{-3}$, vs. $|\frac{\partial t_{term}}{\partial r} \cdot \bar{r}| = 4.12 \times 10^{-2}$, with $\bar{r}$ a characteristic scale for r). Thus changes in t_{term} dominate changes in G as r is varied, and we observe an increase in G. In this case canards lead to local maxima.

The situation is reversed when $t_{term} < t_{term}^{ref}$. As we increase r towards a change in spike number, the last interspike interval and thus t_{term} increase. Because $t_{term} < t_{term}^{ref}$, this increase in t_{term} brings the termination timing closer to the reference termination event, leading to a decrease in G.

These same qualitative features occur if we use global rather than local event detection. The difference is that the landscapes become more jagged, as shown

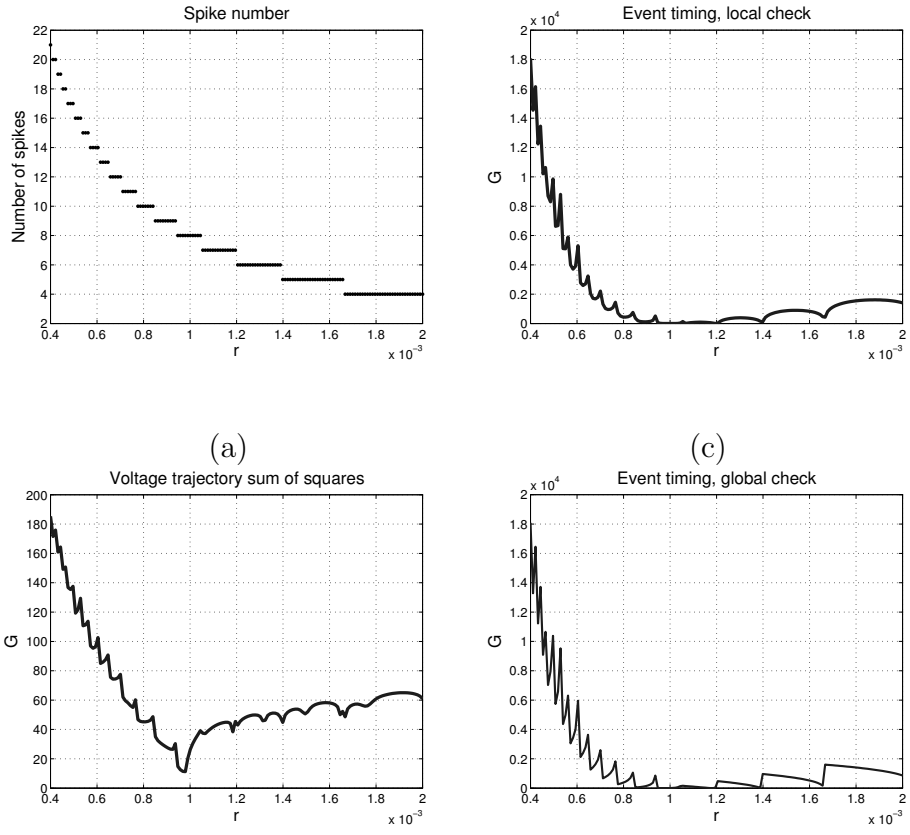


Figure 3.10: Spike number and objective function plots for the Hindmarsh-Rose system (3.5). The reference trajectory corresponds to $(I_{ref}, r_{ref}) = (0.4, 1 \times 10^{-3})$. For all plots $I = 0.39592$, and $a = 1, b = 3, c = 1, d = 5, x_1 = -1.6, S = 4$. (a) Spike number vs. r. (b) Voltage time series squared residuals vs. r. (c) Event timing objective function with local checking vs. r. (d) Event timing objective function with global checking vs. r.

in Figure 3.10d. The global event check considers a termination event valid if there is no spike for $\mathcal{O}(1)$ of slow time following entry into a neighborhood of the critical manifold. Consider the penultimate spike in a burst as we approach a parameter value where the last spike disappears. The "base" of the penultimate spike passes close to the critical manifold. Prior to disappearance of the last spike, the global check determines this close passage of the penultimate spike to be an invalid termination event. As we cross the point where the spike number changes, the last spike disappears. Past this point, there is no longer a spike following the close passage to M_0 and the point becomes a valid termination event. Thus the global event check results in discontinuities in t_{term}.

Based on this discussion, we expect that using a burst event objective function with local event checking will give the best optimization performance. Table 3.1 compares the results from using Levenberg-Marquardt with a line search method for the three different objective functions in trying to fit a transient burst reference trajectory generated from system (3.5) with $(I_{ref}, r_{ref}) = (0.4, 1 \times 10^{-3})$ and initial conditions $(-1.6, -10, 0)$. Five different initial guesses for (I, r) were used. The same reference trajectory corresponding to $(I_{ref}, r_{ref}) = (0.4, 1 \times 10^{-3})$ was used in each case, and the initial phase space values were fixed to match those of the reference trajectory. We see that matching event timing using local checking does produce superior results to the other two objective functions. Two of the five cases converged to the global minimum when using burst events with local checking, whereas none of the cases found the global minimum when using either events with global checking or voltage time series squared residuals. In all cases except for $(I_0, r_0) = (0.5, 5 \times 10^{-3})$, events with local checking resulted in a final spike number closer to the reference spike number than did using the other two methods.

For the $(I_0, r_0) = (0.5, 5 \times 10^{-3})$ case, apparently voltage least squares resulted in the correct number of spikes. However, the final trajectory produced exhibits tonic spiking rather than a transient burst. Thus the final trajectory from voltage least squares is not even qualitatively correct in this case.

Event timing with local checking resulted in excellent fits in two cases $\{(I_0, r_0) = (0.5, 1.5 \times 10^{-3}), (0.35, 8 \times 10^{-4})\}$, reasonable fits in two cases $\{(0.35, 1.8 \times 10^{-3}), (0.65, 7 \times 10^{-4})\}$, and a poor fit in one case $(0.5, 5 \times 10^{-3})$. Fitting the coarse features given by event timing results in improved fits, but the problem of local minima associated with canards persists. The final fit from the starting parameter values of $(0.5, 5 \times 10^{-3})$ is a two spike canard trajectory with a large interspike interval.

In addition to producing the best fits, event timing with local checking also tended to require fewer function evaluations. This should be interpreted with some caution because automatic differentiation was not used for the voltage least squares case. The time mesh used contained 10,000 points. The resulting Jacobians were too large to store in Matlab, and thus derivatives were calculated using finite differences. The increased number of function evaluations may not necessarily mean increased computing time, as it is typically faster to compute a trajectory using a non-Taylor series integrator versus using ADMC++'s Taylor series integrator.

3.2.4 Periodic bursters: fitting period and burst duration

For periodic bursters, burst event timing affects both burst period and duration. These are two qualitative features which are important to capture when trying to fit a model to data. These characteristics are often reported in the literature and

Table 3.1: Optimization output from Levenberg-Marquardt (line-search) algorithm in Matlab's optimization toolbox. The Hindmarsh-Rose system (3.5) was used with active parameters (I, r), with initial active parameter values labeles (I_0, r_0). Reference trajectory corresponded to $(I_{ref}, r_{ref}) = (0.4, 1 \times 10^{-3})$ in each case (see Figure 2.2). The problem type was an IVP with burst events. Automatic differentiation was used for the event timing objective functions.

$(\mathbf{I_0}, \mathbf{r_0})$	**Initial spike number**	**Objective function type**	$(\mathbf{I_{final}}, \mathbf{r_{final}})$	**Final spike number**	**Iterations**	**Function evaluations**
(0.35,1.8$\times 10^{-3}$)	3	burst events (local)	(0.4037,1.229 $\times 10^{-3}$)	7	9	34
		burst events (global)	(0.4339,3.060 $\times 10^{-3}$)	3	34	137
		voltage time series	(0.3076,1.128$\times 10^{-3}$)	5	31	444
(0.5,1.5$\times 10^{-3}$)	7	burst events (local)	(0.4000,1.000 $\times 10^{-3}$)	8	6	22
		burst events (global)	(0.4073,1.447 $\times 10^{-3}$)	6	14	54
		voltage time series	(0.3104,3.986 $\times 10^{-4}$)	16	7	86
(0.35,8$\times 10^{-4}$)	9	burst events (local)	(0.4000,1.000$\times 10^{-3}$)	8	7	29
		burst events (global)	(0.4287,9.319 $\times 10^{-4}$)	9	80	393
		voltage time series	(0.3500,8.000 $\times 10^{-4}$)	9	46	1015
(0.65,7 $\times 10^{-4}$)	21	burst events (local)	(0.6391,1.329 $\times 10^{-3}$)	11	80	345
		burst events (global)	(0.6498,6.982 $\times 10^{-4}$)	21	73	502
		voltage time series	(0.6493,6.976 $\times 10^{-4}$)	22	6	50
(0.5,5 $\times 10^{-3}$)	2	burst events (local)	(0.5025,5.549 $\times 10^{-3}$)	2	80	404
		burst events (global)	(0.4998,8.629 $\times 10^{-3}$)	2	80	323
		voltage time series	(1.856,2.902 $\times 10^{-2}$)	8	26	228

are thus more readily available than complete voltage trace data. Furthermore, fitting the gross features will likely lead to better fits than trying to fit the entire voltage time series, as discussed in section 3.2.3. The question then becomes how best to fit these features. One approach is to first try to fit one characteristic, for example burst period. After a good fit has been obtained, we then try to fit burst duration while maintaining a good burst period fit. Here we consider two methods for this last step: constrained gradient descent, and penalty functions.

Continuation and gradient descent

Section 3.2.2 discussed continuation of periodic orbits with fixed period using a predictor-corrector method. This can be modified easily so that continuation steps are taken in descent directions, for example to match burst duration. Let $\mathcal{M}$ denote the manifold of periodic orbits, and let $\mathcal{M}_T$ denote the codimension 1 submanifold of periodic orbits with period T. Let $m \in \mathcal{M}_T$. The tangent space at m is spanned by the nullspace of $DF_T(m)$ (cf. section 3.2.2). Let $\mathcal{U} = \{u_1, \ldots, u_k\}$ be such a basis for $T_m\mathcal{M}_T$. Note that $k = ap - 1$. As usual, let G denote our objective function. We wish to compute $\nabla G|_{\mathcal{M}_T}$.

Computing $\nabla_\lambda G = D_{\mathbf{\Delta}} G \cdot D_\lambda \mathbf{\Delta}$ is accomplished as discussed in chapter 2. Let $\hat{u}_i$ denote the projection of u_i to the parameter space. Define $\alpha_i = <\nabla_\lambda G, \hat{u}_i>$. Then the steepest descent direction in $\mathcal{M}_T$ is $\nabla G|_{\mathcal{M}_T} = \sum_i \alpha_i u_i$.

We incorporate gradient descent into our continuation framework by taking predictor steps in the $\nabla G|_{\mathcal{M}_T}$ direction. Following the predictor step, we use natural corrector steps [75] to try to locate a descent point on $\mathcal{M}_T$. The step is rejected if the corrector steps fail to converge, or if G does not decrease. In this case we shrink the predictor step and try again. G decreases for sufficiently small

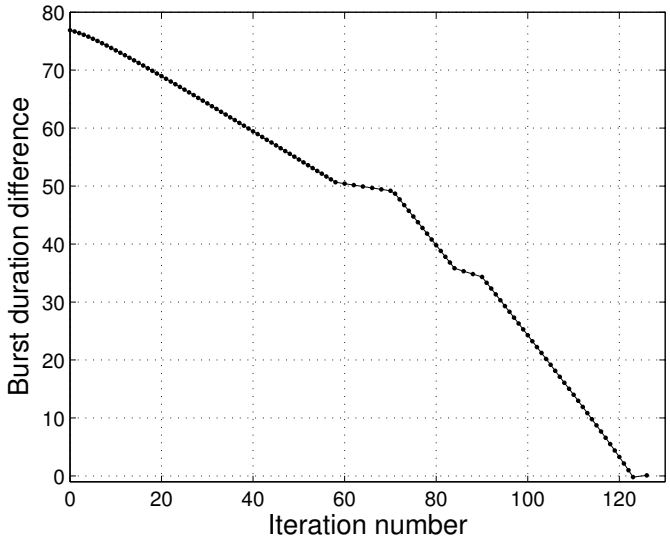


Figure 3.11: Gradient descent along $\mathcal{M}_T$ for the Hindmarsh-Rose system (3.5) with b, I the active parameters.

predictor steps.

As an example problem we take the burst periodic orbit corresponding to system (3.5) with $I = 2, r = 1 \times 10^{-3}$ as our reference trajectory (Figure 3.4), and the burst periodic orbit with $I = 2.2841, r = 1.1938 \times 10^{-3}$ as our initial guess (Figure 3.6). We use gradient descent while keeping the period fixed to that of our initial guess, with active parameters b, I. Figure 3.11 shows the decrease in G at each iteration. The initial discrepancy in burst duration is 76.89. After 116 gradient descent steps, this is reduced to 0.11.

The spike numbers are the same for the initial and final trajectories from the gradient descent algorithm. As observed in section 3.2.2, spike number tends not to change along continuation curves computed using (3.6). During optimization, situations may arise where changing the spike number is desirable. The next section presents one way to allow this.

Penalty functions

It becomes easier for spike number to vary over the course of the optimization algorithm if we relax the condition that the period exactly match a specified value. This can result in improved fits. One way to accomplish this is to use a penalty function which penalizes discrepancies in period. Casey [15] used one such penalty function for fitting non-burst periodic orbits. We use this form of penalty function for both constrained periods and burst durations. Let X denote the burst characteristic we have not yet fit, and Y the burst characteristic that is constrained (e.g. that we have already fit). Then one objective function choice is

$$G = (X - X_{ref})^2 + \left(w \log(\frac{Y_{ref}}{Y})\right)^2. \tag{3.7}$$

This is just the usual squared residual objective function (chapter 2) together with a log penalty term. The penalty weight is given by w. A heuristic for choosing w is that proportional changes in X and Y result in comparable changes in G near (X_{ref}, Y_{ref}). Let $X = (1+\mu)X_{ref}$, $Y = (1+\mu)Y_{ref}$. Then $G = (\mu X_{ref})^2 + \left(w \log(1+\mu)\right)^2$. Setting $\mu X_{ref} = w \log(1+\mu)$ gives $w = \frac{\mu X_{ref}}{\log(1+\mu)} \approx X_{ref}$. Thus $w = X_{ref}$ under this heuristic.

We can use various optimization algorithms for minimizing (3.7), such as the trust region methods discussed in section 1.3. We can also use gradient descent, for example in comparison with the preceding section.

Chapter 4

PreBotzinger neurons

This chapter applies the algorithms developed in chapter 2 to models of respiratory neurons in the preBotzinger complex (pBC). The pBC is a functionally defined region of the brainstem that plays a key role in respiratory rhythm generation [33, 92, 114]. Features of the breathing rhythm such as frequency and tidal volume are modulated in response to environmental conditions such as oxygen [76] and carbon dioxide levels [3], pH [90], temperature [122], and the application of neuromodulators such as norepinephrine [125], serotonin [33, 103], opioids [77, 107], substance P [10, 131], somatostatin [80], dopamine [41, 62, 63], and barbiturates [38]. The pBC includes endogenous bursting cells, termed "pacemakers". Bursting in these pacemaker cells persists in the absence of synaptic connections [32, 73, 92], and burst characteristics for synaptically isolated pacemakers are also subject to modulation [125]. Breathing rhythm disruptions underlie medical conditions such as sleep apnea [8] and Ondine's curse [84], and have been implicated in sudden infant death syndrome (SIDS) [14, 67, 71, 79, 98, 108]. Thus there is high interest in understanding the mechanisms which generate and shape breathing rhythm.

Butera et al [11] presented a model of pBC neurons which contains a minimal set of currents. Viemari and Ramirez [125] measured burst characteristics of synaptically isolated pacemaker cells under different experimental conditions. We use our optimization algorithms to fit model I of Butera et al to the observed burst characteristics. The results illustrate how these algorithms can be used both for model calibration, and for examining different potential modulatory mechanisms.

4.1 Biology of the preBotzinger complex

The pBC is a functionally defined group of cells belonging to a larger region in the brainstem called the ventral respiratory group (VRG). Premotor neurons in the preBotzinger complex project to motor neurons innervating the hypoglossal nerve. Electrical bursts in the hypoglossal nerve correspond to inspiration via movement of the tongue to clear the air passage. Different regions of the VRG are classified according to the timing of their discharge patterns relative to the phrenic nerve, which innervates the diaphragm. Coordination of the different VRG regions is accomplished through excitatory and inhibitory synaptic connections between these regions. However, even when excised from the rest of the VRG, the pBC maintains rhythmic activity similar to the electrical activity of the hypoglossal nerve in the intact system [114]. Lesions to the pBC disrupt the breathing rhythm [33]. Thus the pBC is essential for normal respiratory rhythm generation.

The pBC is not anatomically clearly defined. A proposed definition of the pBC is propriobulbar VRG neurons which express the substance P receptor NK1R [33, 46]. This definition gives an estimate of several hundred neurons in the rat pBC [127]. Excitatory glutamatergic connections exist within the pBC, connecting a variety of cell types including endogenous bursting cells [42, 73]. Understanding how the constituent ion currents of these endogenous bursters influence burst characteristics is of interest, particular as these cells have been proposed to serve as pacemakers for the pBC rhythm [66, 73, 92, 114]. Different types of endogenous bursters possessing different ion channels exist within the pBC [119]. These different bursters may play different functional roles in the breathing rhythm. For example, Cd^{2+}-insensitive bursters may be necessary for "gasping" activity patterns observed under hypoxic conditions [121]. Characterizing the intrinsic membrane

properties of these pacemakers and their modulation is relevant for understanding network state transitions such as the shift from eupnea to gasping [53, 82]. We focus on intrinsic membrane properties and ignore synaptic interactions for the remainder of this chapter.

4.1.1 Ion channels in the preBotzinger complex

An assortment of ion channels are found in the pBC, including fast sodium, delayed rectifier potassium, persistent sodium ($I_{\mathrm{Na(P)}}$), high and low voltage activated calcium [29], nonspecific calcium activated cation channels (I_{CAN}) [119], and hyperpolarization activated cation currents (I_{h}) [118]. Additional inward rectifiers and leak channels are found in the pBC, some of which are considered in further detail below. Various electrical activity patterns are observed in the pBC, including endogenous bursting, tonic spiking, and quiescence. At least two different types of endogenous bursting cells have been identified, those where bursting is abolished in the presence of cadmium, and those which are cadmium insensitive [119]. The former Cd^{2+}-sensitive cells rely upon a I_{CAN} current to generate bursting, whereas the Cd^{2+}-insensitive bursters utilize a persistent sodium current $I_{\mathrm{Na(P)}}$. Additional information on selected pBC currents is given below. We are interested in how the different constituent currents influence burst characteristics such as burst period and duration. Possible modulatory mechanisms are of particular interest, as is understanding the role of the persistent sodium current.

Persistent sodium current. The persistent sodium current $I_{\mathrm{Na(P)}}$ is a voltage-activated sodium current which inactivates on a very slow time scale [18]. $I_{\mathrm{Na(P)}}$ is a widespread current that is found in many mammalian structures. See Crill [18]

and Magistretti and Alonso [81] for listings. As reviewed by Magistretti and Alonso [81], $I_{\mathrm{Na(P)}}$ plays many functional roles including burst generation (such as in the pBC), bistability and plateau potential generation, and excitatory postsynaptic potential amplification [21]. The persistent sodium current has been implicated in several medical conditions, including hyperkalemic periodic paralysis [13] and certain forms of epilepsy [110].

Despite its importance, the persistent sodium current is poorly characterized in many systems. Kinetic data is incomplete and mechanisms for persistence are not understood. It is unknown whether $I_{\mathrm{Na(P)}}$ channels comprise a distinct family of channels from the fast sodium channels, if $I_{\mathrm{Na(P)}}$ channels are a splice variant of other fast channels, or if some other mechanism is involved. Some examples of kinetic data for $I_{\mathrm{Na(P)}}$ come from the hippocampus [39, 40] and the entorhinal cortex [81]. Modulation of $I_{\mathrm{Na(P)}}$ is largely unexplored. Thus modeling can play an important role in elucidating possible roles for the persistent sodium current.

Calcium currents. Both high voltage activated (HVA) and low voltage activated (LVA) calcium currents are expressed in the pBC [29]. Calcium imaging has been used to identify bursting neurons in the pBC [73]. These calcium currents activate Cd^{2+}-sensitive pacemakers, and may be important in development [30]. Johnson et al [66] demonstrated that rhythmic activity in the pBC persists even in low calcium media. Because of this, calcium currents are omitted from Butera et al's model and we will not consider calcium currents here.

Calcium activated nonspecific cation current. Two pacemaker types have been found in the pBC. One relies upon the calcium activated nonspecific cation current for burst generation, and the other relies upon the persistent sodium cur-

rent. In the presence of Cd^{2+}, the I_{CAN}-dependent burster ceases bursting, while bursting persists in the $I_{Na(P)}$-dependent burster. Hence these burster types are called Cd^{2+}-sensitive and Cd^{2+}-insensitive bursters, respectively. Cd^{2+}-sensitive and Cd^{2+}-insensitive bursters respond differently to neuromodulators. Specifically, Viemari and Ramirez [125] found different responses of synaptically isolated pacemakers to norepinephrine. Norepinephrine caused burst duration to increase in Cd^{2+}-sensitive pacemakers, while burst period was not changed. For Cd^{2+}-insensitive pacemakers, burst period decreased but burst duration did not alter significantly following norepinephrine application.

Butera et al's model [11] omits Ca^{2+} currents, and hence calcium-activated currents. We focus on fitting Butera et al's model to data for Cd^{2+}-insensitive pacemakers.

Inward rectifiers, sags, and leaks. There are a number of additional channels present in the pBC, including the hyperpolarization activated cation current I_h [118], inward rectifier potassium currents such as ATP-sensitive potassium channels [52, 85, 86], and the pH-sensitive TASK potassium channels [129, 130]. We group these currents together because their conductances are all dominated by potassium, and thus their reversal potentials are close both to one another and to the resting potential for pBC cells. This suggests modeling these currents by approximating them with an effective leak current (section 3.1).

4.1.2 Modulation of the respiratory rhythm

Breathing patterns change in response to environmental variables such as oxygen and carbon dioxide partial pressures, pH, temperature, and a wide range of

neuromodulators including norepinephrine, serotonin, substance P, dopamine, somatostatin, and opioids. Effects include altered breath frequency, tidal volume, and induced apnea. Both synaptic interactions and intrinsic membrane properties can be modulated, and short and long term effects are both observed [33, 87]. Knowledge of these modulatory mechanisms is incomplete. There is a large literature on respiratory rhythm modulation. We mention below a few topics from the literature that will be most relevant for our optimization work on bursting in Butera et al's model.

Norepinephrine. Norepinephrine is a respiratory system neuromodulator that plays an important role in responses to elevated carbon dioxide levels [70]. Viemari and Ramirez [125] examined the effect of norepinephrine on synaptically isolated pBC pacemakers, and found different effects on Cd^{2+}-sensitive and Cd^{2+}-insensitive pacemakers. Norepinephrine altered Cd^{2+}-insensitive pacemaker burst period, but did not affect burst duration. In contrast, Cd^{2+}-sensitive pacemaker burst duration was modulated but burst period was unaffected. Both α_1 and α_2 noradrenergic receptors affect respiratory output through G protein coupled pathways [55, 66, 125]. Norepinephrine acts on synaptically isolated Cd^{2+}-sensitive pacemakers through an α_1-receptor mediated effect on calcium dependent channels [125]. The modulatory target for Cd^{2+}-insensitive pacemakers is not known.

Modulation of inward rectifiers, sags, and leaks. Many respiratory neuromodulators act by regulating K^+ channel activity, often through G protein coupled pathways. Serotonin has myriad effects on respiration. The binding of serotonin to 5-HT1A receptors activates inward rectifier potassium channels K_{ir} [103]. Both opioids and GABA have also been conjectured to regulate K_{ir} activity in the pBC

via $G_{i/o}$ pathways [65]. One specific type of inward rectifier is the ATP-sensitive potassium channel K_{ATP}. K_{ATP} channels in the plasmalemma are involved in short term respiratory depression following hypoxia [52]. Mitochondrial K_{ATP} channels also exist. These are conjectured to play a role in long term facilitation following repeated hypoxic episodes [85].

Serotonin binding of 5-HT4 receptors also affects I_{h} activity via cAMP-mediated phosphorylation [7]. Brain derived neurotrophic factor (BDNF) also modulates I_{h}. Thoby-Brisson et al [118] showed that BDNF application both decreased I_{h} conductance and shifted I_{h} activation curves to more hyperpolarized voltages in neonatal mice.

The pH-sensitive TASK channels are another group of K^+ channels which are modulatory targets in the pBC [129, 130]. TASK-1 and TASK-3 channels are expressed in the rat pBC [129]. These channels are sensitive to pH and to inhalation anesthetics such as halothane [129], both of which influence the breathing rhythm.

4.2 The model of Butera, Rinzel and Smith

As discussed in section 4.1.1, not all of the currents found in the pBC are necessary for endogenous bursting. Butera et al [11] developed a Hodgkin-Huxley style model for pBC cells which includes a minimal set of currents: a fast sodium current I_{Na}, a delayed rectifier potassium current I_{K}, and a persistent sodium current $I_{\mathrm{Na(P)}}$. The model equations are:

$$
\begin{aligned}
C\dot{v} &= -I_{\mathrm{Na}} - I_{\mathrm{K}} - I_{\mathrm{Na(P)}} - I_L && (4.1)\\
I_{\mathrm{Na}} &= g_{\mathrm{Na}} m_\infty^3 (1-n)(v - v_{\mathrm{Na}}) && (4.2)\\
I_{\mathrm{K}} &= g_{\mathrm{K}} n^4 (v - v_{\mathrm{K}}) && (4.3)\\
I_{\mathrm{Na(P)}} &= g_{\mathrm{Na(P)}} \hat{m}_\infty (v - v_{\mathrm{Na(P)}}) && (4.4)\\
I_L &= g_{\mathrm{L}} (v - v_{\mathrm{L}}). && (4.5)
\end{aligned}
$$

The membrane potential is given by v, and $m, n, \hat{m}$, and h are gating variables (Table 4.1). m and n are activation variables for the fast sodium and delayed rectifier potassium channels, while $\hat{m}$ and h are the activation and inactivation variables for the persistent sodium channel. The variable n also serves as the inactivation variable for the fast sodium channel. Equations for the gating variables share a common form:

$$
\begin{aligned}
\dot{x} &= \frac{x_\infty - x}{\tau_x} && (4.6)\\
x_\infty(v) &= \frac{1}{1 + \exp(\frac{v-\theta_x}{\sigma_x})} && (4.7)\\
\tau_x(v) &= \frac{\bar{\tau}_x}{\cosh(\frac{v-\theta_x}{2\sigma_x})}. && (4.8)
\end{aligned}
$$

Each gating variable is associated with three parameters, $\theta, \bar{\tau}$, and σ. x_∞ denotes the steady state value for gating variable x when the voltage is held fixed. θ is the half-activation value ($x_\infty(\theta) = \frac{1}{2}$), and σ gives the steepness of the steady state activation curve. The sign of σ indicates whether the steady state value x_∞ increases (negative σ) or decreases with voltage. For a fixed voltage, gating variables approach their steady state values x_∞ exponentially at a rate $\tau(v)$. The value of $\tau(v)$ is given by a bell-shaped curve with peak of height $\bar{\tau}$ centered at θ

Table 4.1: Phase space variables for system (4.1)-(4.8).

v	membrane potential (mV)
t	time (ms)
m	fast sodium activation
n	potassium activation and fast sodium inactivation
$\hat{m}$	persistent sodium activation
h	persistent sodium inactivation

and half-width given by σ.

Table 4.2 lists the units and default values of the parameters for system (4.1)-(4.8). Butera et al [11] also include an external applied current and a tonic excitatory current. We omit these because they can be absorbed into the leak current, as discussed in section 3.1.

System (4.1)-(4.8) involves several simplifying assumptions. Activation of both the fast and persistent sodium currents is assumed to be instantaneous. Additionally, fast sodium current inactivation is assumed to be linearly related to activation of the potassium current [35]. These are all standard assumptions.

4.3 Optimization for Butera et al's model

Viemari and Ramirez [125] measured burst duration and period for synaptically isolated pBC pacemaker cells. Here we fit system (4.1)-(4.8) to match burst characteristics measured for Cd^{2+}-insensitive bursters both in control situations and after norepinephrine application. The empirical data is given in Table 4.3. Throughout this chapter we use the Levenberg-Marquardt optimization algorithm, together

Table 4.2: Parameters for system (4.1)-(4.8) [11].

C	membrane capacitance	21 pF
g_{Na}	maximal fast sodium conductance	28 nS
v_{Na}	fast sodium reversal potential	50 mV
θ_m	m half-activation	-34 mV
σ_m	m time scale half-width	-5 mV
g_{K}	maximal potassium conductance	11.2 nS
v_{K}	potassium reversal potential	-85 mV
θ_n	n half-activation	-29 mV
σ_n	n time scale half-width	-4 mV
$\bar{\tau}_n$	n time scale constant	10
$g_{\mathrm{Na(P)}}$	maximal persistent sodium conductance	2.8 nS
$v_{\mathrm{Na(P)}}$	persistent sodium current reversal potential	50 mV
$\theta_{\hat{m}}$	$\hat{m}$ half-activation	-40 mV
$\sigma_{\hat{m}}$	$\hat{m}$ time scale half-width	-6 mV
θ_h	h half-activation	-48 mV
σ_h	h time scale half-width	6 mV
$\bar{\tau}_h$	h time scale constant	1×10^4
g_{L}	maximal leak conductance	2.8 nS
v_{L}	leak reversal potential	-60 mV

Table 4.3: Empirical data for Cd^{2+}-insensitive pacemakers. From Table 4 of Viemari and Ramirez [125].

	Control	Norepinephrine
Burst duration (s)	$0.44 \pm .02$	0.52 ± 0.05
Frequency (Hz)	0.22 ± 0.05	0.41 ± 0.06
Mean Period (ms)	4546	2439

with various objective functions.

4.3.1 Control data fits

We first seek to match the burst characteristics reported for Cd^{2+}-insensitive pacemakers in the control setting (see Table 4.3). This is accomplished by first matching burst period, and then matching burst duration while maintaining a good fit for the period. The leak parameters affect burst period [11] and are a likely target of modulation (section 4.1.2). Thus we use $(g_{\mathrm{L}}, v_{\mathrm{L}})$ as our initial active optimization parameters. The observed mean control burst period was 4546 ms and burst duration averaged 440 ms. We started with an initial burster at $(g_{\mathrm{L}}, v_{\mathrm{L}}) = (2.8, -59.3)$ with period 4308 ms and burst duration 797 ms. Applying our optimization algorithms with (2.15) as the objective function resulted in the parameter values labeled Set 1 in Table 4.4. The period matches the empirical observations very closely, but the burst duration of 749 ms is longer than the actual 440 ms observed in the control conditions. The Set 1 voltage trace is shown in Figure 4.1b.

We next start from Set 1 and use a log penalty function as described in section 3.2.4 to try to fit the burst duration while maintaining a good fit for the period. There are again many choices for active optimization parameters. The period and

Table 4.4: Two parameter sets for system (4.1)-(4.8) obtained from optimization. Set 1 resulted from attempting to match the period of Cd^{2+}-insensitive pacemakers in the control setting (Table 4.3), with active parameters g_L, v_L. Control 1 the output from attempting to match both period and burst duration for Cd^{2+}-insensitive pacemakers in the control setting, starting from Set 1 with active parameters $\theta_m, \theta_n, \sigma_n, \bar{\tau}_n, \sigma_h$. All other parameter values given in Table 4.2.

	Active Optimization Parameters		Period (ms)	Burst duration (ms)
Set 1	$g_L = 2.83131935965688$	$v_L = -59.30989949043865$	4546	769
Control 1	$\theta_m = -34.11588874462502$	$\theta_n - 27.23254555031786$	4548	440
	$\sigma_n = -4.38901615009862$	$\bar{\tau}_n = 9.39771256441101$		
	$\sigma_h = 5.17827437646577$			

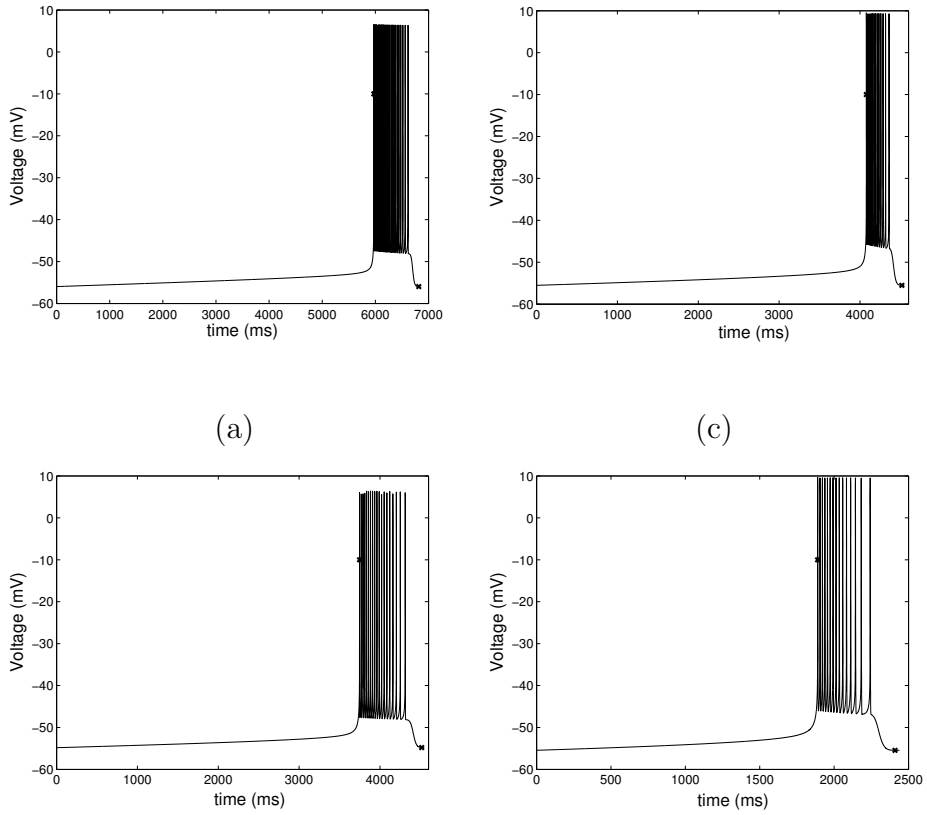


Figure 4.1: Four bursting periodic orbits. (a) Default parameter values, Table 4.2. (b) Set 1, Table 4.4. (c) Control 1, Table 4.4. (d) NE 1, Table 4.8.

burst duration sensitivities can aid in active parameter selection. Table 4.5 lists the scaled sensitivities computed using automatic differentiation, together with their rankings by magnitude. The rankings give the index after sorting by magnitude in descending order. We say that a parameter has a high period sensitivity ranking if the period sensitivity to this parameter is large compared with other period sensitivities. In other words, high ranking sensitivities correspond to small indices.

We are trying to adjust the burst duration while keeping the period unchanged. Thus a heuristic is to look for parameters which affect burst duration more strongly than period, which means using parameters with higher burst rank than period

Table 4.5: Parameter sensitivities for the Set 1 burster (Table 4.4). Scale factors correspond to the parameter values in Table 4.2.

Parameter	Scaled sensitivities		T rank	Burst rank
	$\frac{\partial T}{\partial \lambda}$	$\frac{\partial (t_{term} - tinit)}{\partial \lambda}$		
g_{Na}	-8.11×10^2	-6.16×10^2	17	17
v_{Na}	-8.09×10^2	-7.19×10^2	18	16
θ_m	1.42×10^3	-2.21×10^3	15	10
σ_m	5.61×10^2	-8.03×10^2	19	15
g_{K}	-2.20×10^3	-1.28×10^3	13	11
v_{K}	4.83×10^3	3.53×10^3	9	7
θ_n	1.18×10^4	6.87×10^3	7	3
σ_n	1.33×10^3	5.10×10^3	16	5
$\bar{\tau}_n$	-1.88×10^3	-2.67×10^3	14	8
$g_{\mathrm{Na(P)}}$	-2.15×10^4	9.44×10^2	5	14
$v_{\mathrm{Na(P)}}$	-1.07×10^4	5.92×10^2	8	18
$\theta_{\hat{m}}$	1.26×10^5	-10.62	2	19
$\sigma_{\hat{m}}$	3.54×10^4	5.79×10^3	4	4
θ_h	-5.96×10^4	8.05×10^3	3	2
σ_h	3.59×10^3	4.01×10^3	11	6
$\bar{\tau}_h$	2.79×10^3	1.24×10^3	12	13
g_{L}	2.07×10^4	-1.24×10^3	6	12
v_{L}	-1.31×10^5	-2.27×10^4	1	1
C	3.63×10^3	2.21×10^3	10	9

rank. The parameters $\theta_m, \sigma_m, \theta_n, \sigma_n, \bar{\tau}_n$, and σ_h all fit this description. We discard σ_m because its burst sensitivity ranking is low. Using the remaining five active parameters with objective function (3.7) results in the parameter set labeled as Control 1 in Table 4.4. Essentially a perfect fit is obtained. The voltage trace is shown in Figure 4.1c.

Another approach is to select active parameters based upon biological motivations. The persistent sodium current is incompletely characterized. As the persistent sodium current parameter values in system (4.1)-(4.8) are uncertain, it is sensible to search for better fits to data by modifying these parameters. The parameter set labeled Control 2 in Table 4.6 shows the results of optimizing over all the $I_{\mathrm{Na(P)}}$ parameters. Again a perfect fit is obtained.

In both Control 1 and 2 the relative change in σ_h over the course of the optimization is greater than that for the other active parameters, suggesting that σ_h is important for determining burst characteristics. We examine this further by starting from Set 1 and repeating the optimization using a smaller subset of the $I_{\mathrm{Na(P)}}$ parameters which includes σ_h. Control 5 shows the results using (θ_h, σ_h), and Control 6 the results from $(g_{\mathrm{Na(P)}}, \sigma_h)$. The resulting fits are substantial improvements from the initial burster corresponding to Set 1: burst duration discrepancies are within 10% for Control 5 and 18% for Control 6, while the period remains within 1% of the empirical observations. Burst duration in the presence of norepinephrine was 520 ms, which was not significantly different from the control burst duration of 440 ms (Table 4.3,[125]). The fits from Controls 5 and 6 both deviate less than this from the control burst duration.

Table 4.5 shows that the relative influence of the potassium current parameters is greater on the burst duration than on the period. Control 4 in Table 4.6 shows

the results of including $\bar{\tau}_n$ together with σ_h in the active optimization parameters. Both $\bar{\tau}_n$ and σ_h change significantly over the course of the optimization; the resulting fit is excellent.

Neuromodulators often act by altering conductances. We investigate this possibility by optimizing over all conductances. The resulting conductance values are given in Control 3 of Table 4.6. A perfect fit is obtained. The slow conductances $g_{\mathrm{Na(P)}}$ and g_{L} experience the greatest relative change over the course of the optimization. The potassium conductance g_{K} also experiences a large relative change of 45%, whereas the fast sodium conductance changes by less than 4%.

Leak channel modulation is one of the primary known modulatory mechanisms in the pBC (4.1.2). However, using $(g_{\mathrm{L}}, v_{\mathrm{L}})$ as active parameters to try to fit the control data from Set 1 results in a poor fit, labeled Control 8 in Table 4.6. The final burst duration is 638 ms, much greater than the desired 440 ms. Of the fits reported in Table 4.6, the Control 8 period is also the farthest from the mean observed control period.

4.3.2 Norepinephrine data fits

Norepinephrine application to Cd^{2+}-insensitive pacemakers causes burst period to shorten, but does not significantly alter burst duration (Table 4.3, [125]). We examine different mechanisms for this change by starting from one of the obtained fits to the control data discussed in section 4.3.1, and using our optimization algorithms to match the burst characteristics in the norepinephrine condition. Parameter sets Control 1-4 in Tables 4.4 and 4.6 are all excellent fits to the control data. We use Control 1 as the starting point for optimization, but could just as easily have used Controls 2-4 as well.

Table 4.6: Obtained fits for Cd^{2+}-insensitive pacemakers in the control setting (Table 4.3). Initial parameters correspond to Set 1 in Table 4.4.

	Active Optimization Parameters		Period (ms)	Burst duration (ms)
Control 2	$g_{\mathrm{Na(P)}} = 2.78803635332670$	$v_{\mathrm{Na(P)}} = 50.00038456341981$	4546	440
	$\theta_{\hat{m}} = -39.90657397120774$	$\sigma_{\hat{m}} = -5.86441933703251$		
	$\theta_h = -48.18443018606493$	$\sigma_h = 4.80494852947231$		
	$\bar{\tau}_h = 9.999999797716771 \times 10^3$			
Control 3	$g_{\mathrm{Na}} = 29.03862300309391$	$g_{\mathrm{K}} = 16.20959468054647$	4546	440
	$g_{\mathrm{Na(P)}} = 5.31769065347834$	$g_{\mathrm{L}} = 5.70345313826121$		
Control 4	$\tau_n = 6.18021288547017$	$\sigma_h = 4.27165795051380$	4466	446
Control 5	$\theta_h = -48.91041735798517$	$\sigma_h = 5.15141990125337$	4531	482
Control 6	$g_{\mathrm{Na(P)}} = 2.65391538068674$	$\sigma_h = 5.15791458196960$	4516	516
Control 7	$g_{\mathrm{Na(P)}} = 6.96795374561904$	$g_{\mathrm{L}} = 6.88180511304864$	4584	483
Control 8	$g_{\mathrm{L}} = 3.15102944281655$	$v_{\mathrm{L}} = -58.57622465082586$	4952	638

Table 4.7 gives the burst characteristic sensitivities for the Control 1 burster. The highest period sensitivity rankings occur for slow current ($I_{\mathrm{Na(P)}}, I_L$) parameters. The parameters which have higher period sensitivity rankings than burst duration sensitivity rankings are $g_{\mathrm{Na(P)}}, v_{\mathrm{Na(P)}}, \theta_{\hat{m}}, \theta_h, \sigma_h, \bar{\tau}_h$, and g_{L}. As conductances are a common modulatory target, we select $g_{\mathrm{Na(P)}}$ and g_{L} from this set for our optimization parameters. The resulting output is labeled as NE 1 in Table 4.8, and the voltage trace is shown in Figure 4.1d. A perfect fit to the observed period and burst duration is obtained. This importance of the slow conductances is also seen if we allow all conductances to vary, as shown in NE 2 of Table 4.8. A perfect fit is again obtained, with the largest relative changes in the active parameter values occurring for $g_{\mathrm{Na(P)}}$ and g_{L}. Note that the relative change in g_{K} is small. This is in contrast to the fit for burst duration with constrained period obtained in Control 3 (Table 4.6). Varying the leak parameters $g_{\mathrm{L}}, v_{\mathrm{L}}$ is another way to match the observed burst characteristics in the norepinephrine condition. The parameter set NE 3 in Table 4.8 shows the final parameter values obtained using $g_{\mathrm{L}}, v_{\mathrm{L}}$ as the active parameters. An exact match is obtained, in contrast with attempts to fit the control data using the leak parameters (Control 8). On the other hand, varying (τ_n, σ_h) resulted in an excellent fit to the control data. Attempting to fit the norepinephrine data using (τ_n, σ_h) results in a poor fit (NE 7).

We can also attempt to use the persistent sodium current to match the norepinephrine data. Perfect fits are obtained when using $g_{\mathrm{Na(P)}}, \theta_{\hat{m}}, \sigma_{\hat{m}}, \theta_h, \sigma_h, \bar{\tau}_h$ as active parameters, as shown in NE 4, Table 4.8. An exact match is still obtained if we reduce the active parameter set to $(g_{\mathrm{Na(P)}}, \sigma_h)$, corresponding to NE 5, or to (θ_h, σ_h), labeled NE 6 in Table 4.8.

Table 4.7: Parameter sensitivities for the Control 1 burster (Table 4.4). Scale factors correspond to the parameter values in Table 4.2.

Parameter	Scaled sensitivities		T rank	Burst rank
	$\frac{\partial T}{\partial \lambda}$	$\frac{\partial (t_{term} - tinit)}{\partial \lambda}$		
g_{Na}	5.22×10^2	1.31×10^2	17	16
v_{Na}	7.56×10^2	2.23×10^2	16	15
θ_m	5.06×10^3	1.99×10^3	10	3
σ_m	1.66×10^3	4.41×10^2	12	10
g_{K}	-80.36	5.54×10^2	19	5
v_{K}	-9.23×10^2	-1.31×10^3	15	6
θ_n	1.37×10^3	-2.98×10^3	14	2
σ_n	-6.44×10^3	-8.45×10^2	8	7
$\bar{\tau}_n$	2.82×10^3	8.10×10^2	11	8
$g_{\mathrm{Na(P)}}$	-1.77×10^4	-63.43	5	17
$v_{\mathrm{Na(P)}}$	-8.82×10^3	-62.48	7	18
$\theta_{\hat{m}}$	9.71×10^4	-1.61×10^3	1	5
$\sigma_{\hat{m}}$	2.05×10^4	-1.71×10^3	4	4
θ_h	-6.19×10^4	-4.36×10^2	3	11
σ_h	-5.71×10^3	-2.69×10^2	9	14
$\bar{\tau}_h$	1.59×10^3	-40.62	13	19
g_{L}	1.68×10^4	-3.32×10^2	6	12
v_{L}	-6.96×10^4	6.35×10^3	2	1
C	3.03×10^2	-2.86×10^2	18	13

Table 4.8: Obtained fits for Cd^{2+}-insensitive pacemakers with norepinephrine (Table 4.3). Initial parameters correspond to Control 1 in Table 4.4.

	Active Optimization Parameters		Period (ms)	Burst duration (ms)
NE 1	$g_{\mathrm{Na(P)}} = 3.13886124846613$	$g_{\mathrm{L}} = 2.45042430012711$	2439	520
NE 2	$g_{\mathrm{Na}} = 28.02025609502303$	$g_{\mathrm{K}} = 11.37248860550097$	2439	520
	$g_{\mathrm{Na(P)}} = 2.92820956943392$	$g_{\mathrm{L}} = 2.33170555842205$		
NE 3	$g_{\mathrm{L}} = 2.41245208422888$	$v_{\mathrm{L}} = -58.71360221032916$	2439	520
NE 4	$g_{\mathrm{Na(P)}} = 3.00550132155538$	$\theta_{\hat{m}} = -40.12438473300712$	2439	520
	$\sigma_{\hat{m}} = -6.38932351933371$	$\theta_h = -47.96424048323226$		
	$\sigma_h = 5.15967228260065$	$\bar{\tau}_h = 9.999999989757554 \times 10^3$		
NE 5	$g_{\mathrm{Na(P)}} = 3.60948821726906$	$\sigma_h = 5.28349994032101$	2439	520
NE 6	$\theta_h = -44.33010676170945$	$\sigma_h = 4.73543970700419$	2439	520
NE 7	$\bar{\tau}_n = 9.85382827997729$	$\sigma_h = 5.37770930433051$	4528	494

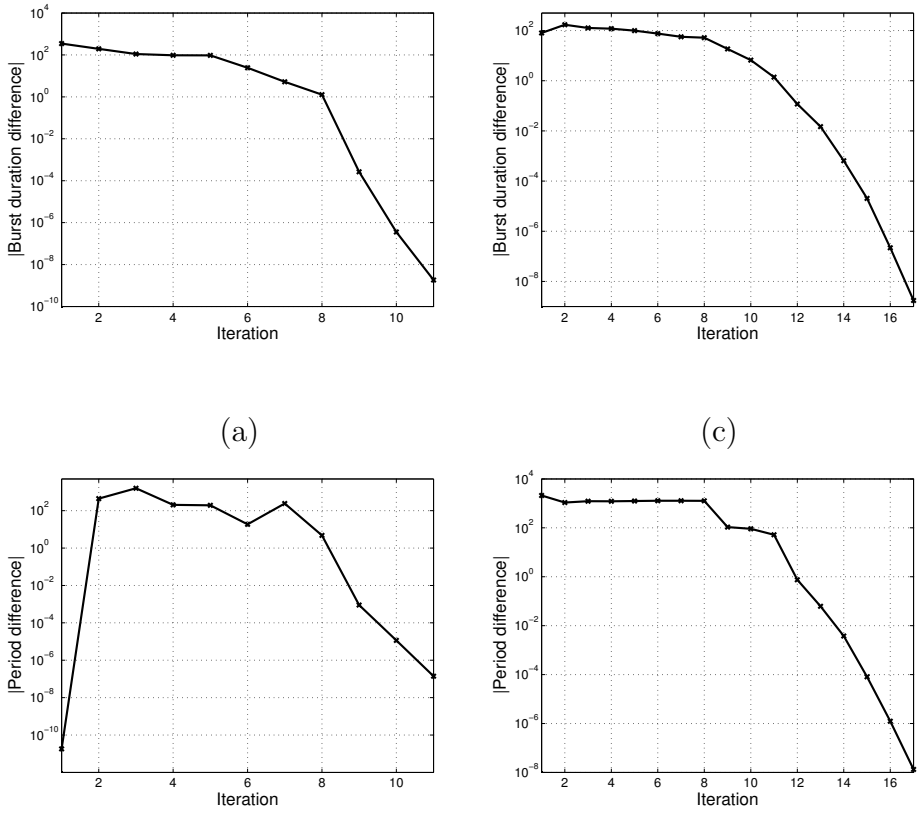


Figure 4.2: Progress of two optimization runs. (a)-(b): Optimization run resulting in Control 2, Table 4.6. Differences in period were penalized. (c)-(d): Optimization run resulting in NE 6, Table 4.8. Differences in burst duration were penalized.

4.4 Discussion

Burst features such as period, burst duration, and number of action potentials per burst are all of biological interest, determining features such as secretory rates and the frequency and force of muscle contraction [43]. Our investigations show that system (4.1)-(4.8) is flexible, and can control these features independently of one another by many different means. Bursting involves complicated dynamics. As observed in section 3.2, spike number, burst duration, and burst period are all linked with one another, with transitions in spike number bringing with it

concomitant changes in period. These transitions involve delicate phenomena such as canards. Perhaps because of this complexity, few studies have attempted to estimate parameters for bursting models. The studies which do exist either use non-smooth methods (for example, using objective functions which depend explicitly on spike number [6, 124]), or are brute-force methods involving simulations on multi-dimensional grids in parameter space to assemble large databases, which are subsequently searched to find best fits to desired characteristics [99]. Few if any studies have been published which estimate bursting model parameters based upon any sort of local information. This study appears novel in this regard. Hopefully it is also useful: good fits can be obtained, and local information can be informative, as can information from the progress of the optimization algorithms.

Optimization algorithms can be useful for model calibration. Butera et al [11] developed and selected parameter values for system (4.1)-(4.8) based on a mix of empirical measurements and computer simulations. Existence of the persistent sodium current in the pBC had not yet been established. There is no published record of attempts to fit model output quantitatively to voltage recordings. This is reasonable, as the model is a minimal one intended to capture qualitative behavior rather than to fit voltage traces precisely. The work here shows that smooth optimization methods can be used to better understand these qualitative features. Introducing defining equations for burst initiation and termination allows the use of smooth optimization methods to quantitatively fit these burst characteristics. This can aid in model calibration, help generate hypotheses for biological modulatory mechanisms, and facilitate exploration of model behavior.

Using the control data of Viemari and Ramirez [125] to calibrate system (4.1)-(4.8) shows that good fits can be obtained in a variety of ways. Excellent fits can be

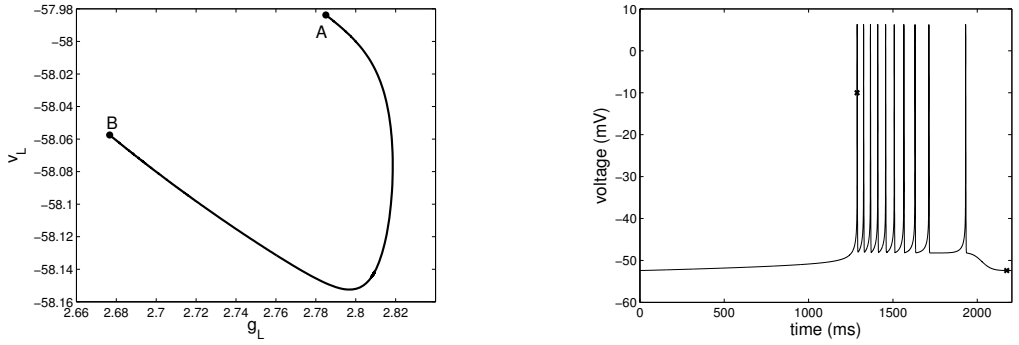
obtained by adjusting a collection of activation and time constant parameters from all the currents (Control 1, Table 4.4), all parameters of the persistent sodium current (Control 2, Table 4.6), time constant and activation parameters for potassium and persistent sodium inactivation (Control 4), and conductances for all currents in the model (Control 3). It is known that model cells have many ways to generate similar output [37, 44]. Perhaps this should be expected: bursting is a fundamental activity pattern that is essential for numerous biological processes, so having redundancy in the system is sensible. What is useful from the parameter fitting done here is the insight gained for the relative importance of different parameters in obtaining good fits. For example, in Control 2 we vary 7 parameters but see that the relative change in σ_h is by far the largest over the course of the optimization. This suggests that σ_h is important for altering burst duration when the period is constrained. This is supported by subsequent optimization runs which include σ_h in a smaller set of parameters (e.g. Control 4), where an excellent fit is still obtained. Model calibration is of course also of interest in and of itself. Having more realistic models allows the testing of more specific questions and gives more confidence in model relevance.

That there are many ways for cells to fire similar burst patterns can be seen through the period and burst duration sensitivities. Let $D_\lambda T$ denote the Jacobian of the period with respect to parameters. The entries of $D_\lambda T$ consist of the period sensitivities. The kernel of $D_\lambda T$ gives the parameter tangent space of $\mathcal{M}_T$, the manifold of periodic orbits with fixed period T. This information can be used in a predictor-corrector framework for continuation, as discussed in section 3.2.2. Figure 4.3 shows an example for continuation of fixed period orbits in the $(g_{\mathrm{L}}, v_{\mathrm{L}})$ plane. Comparing the trajectories from points A and B in Figure 4.3a, we see that

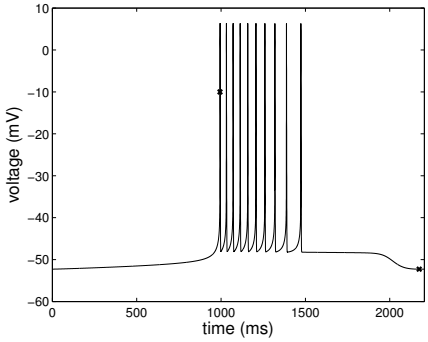
burst duration changes significantly over $\mathcal{M}_T$ (Figures 4.3b-c). Similarly, the kernel of $D_\lambda(t_{term} - t_{init})$ gives the parameter tangent space of the manifold of periodic orbits with constant burst duration. The intersection of these two manifolds gives the submanifold of fixed period and burst duration orbits.

Predictor-corrector continuation tends to follow burst families with a fixed number of spikes (Figure 4.3, cf. section 3.2.2). On the other hand, spike number does change when using penalty functions to fit the control data. For example, Set 1 has 19 spikes while Control 1 has 14 (Figure 4.1). Figure 4.2 shows discrepancies in burst duration and period over the course of the optimization runs which terminate at Control 2 and NE 6. For Control 2 discrepancies in period are penalized, while the NE 6 optimization run uses a burst duration penalty. In both cases we see that the penalty initially increases before eventually decreasing. This initial increase allows for changes in spike number.

One of our objectives in this study was to examine possible roles of the persistent sodium current in shaping bursts. From Control 2 we see that $I_{\mathrm{Na(P)}}$ can modify burst duration without changing burst period. The inactivation kinetics appear to be particularly important. The persistent sodium current can also be used to modulate burst period while the burst duration is constrained. This is seen in NE 4, where a perfect fit to the norepinephrine data is obtained by modulating $I_{\mathrm{Na(P)}}$ parameters. However, here the largest relative changes occur for $g_{\mathrm{Na(P)}}$ and $\theta_{\hat{m}}$, whereas σ_h changes much less. The results for NE 1,2, and 4 in Table 4.8 suggest that changes in the slow conductances $g_{\mathrm{Na(P)}}, g_{\mathrm{L}}$ are a ready means for modulating period independent of burst duration. This is supported by the sensitivities in Table 4.7. The parameters $g_{\mathrm{Na(P)}}, v_{\mathrm{Na(P)}}, \theta_{\hat{m}}, \theta_h, \sigma_h, \bar{\tau}_h$,and g_{L} all have higher ranking period sensitivities than burst duration sensitivities. Within this



(a) (b)



(c)

Figure 4.3: Continuation of fixed period bursters for system (4.1)-(4.8) in the $(g_{\mathrm{L}}, v_{\mathrm{L}})$ plane. (a) The manifold $\mathcal{M}_T$ of fixed period bursters. (b) Burst orbit at point A of $\mathcal{M}_T$. (c) Burst orbit at point B of $\mathcal{M}_T$.

set, the period is least sensitive to σ_h and $\bar{\tau}_h$. For both Set 1 and Control 1, the slow conductances have period sensitivity rankings in the top third (Tables 4.5-4.7). Thus slow conductances are a prime candidate for period modulation when there are burst length constraints.

On the other hand, for both Set 1 and Control 1 the potassium current parameters have higher burst sensitivity rankings than period sensitivity rankings. This is consistent with Ghigliazza and Holmes [43], who found that changing the potassium current time scale in the SRK model [113] could strongly affect the number of spikes per burst while only moderately affecting burst period. The fast sodium activation parameters θ_m, σ_m also appear to affect burst length more strongly than period. The period and burst duration are both very sensitive to the leak reversal potential. This is true for both Set 1 and Control 1. A perfect fit to the burst characteristics in the presence of norepinephrine can be obtained by varying the leak parameters $g_{\mathrm{L}}, v_{\mathrm{L}}$ (NE 3). Thus varying the leak parameters is one likely modulatory mechanism for altering burst period with minimal disturbance to burst duration. It may be more difficult to use the leak parameters to alter burst duration while respecting constraints on the period (see Control 8). Likely candidates for modulating burst duration independently of period include the delayed rectifier potassium currents (currents which operate on the middle time scale), and time scale parameters for slow processes such as persistent sodium current inactivation.

Chapter 5

Conclusions and future work

This thesis uses smooth local optimization methods which are based on the structure of the ODEs. Specifically, we take advantage of the geometry underlying bursting by introducing burst event defining equations, and use these defining equations together with automatic differentiation to fit burst characteristics. This approach is new, combining optimization methods with fast-slow decompositions of bursting trajectories of model equations.

Smooth local methods. Gradient-based local optimization methods are mathematically well behaved and tractable for analysis. They are less computationally intensive than global methods such as simulated annealing or stochastic search, and do not have the difficulties inherent in brute force simulations through high dimensional grids of parameter space. This facilitates exploring changes in model structure. Local information also has inherent biological interest. For example, the period and burst timing parameter sensitivities help tease apart the roles of different ion currents in shaping bursts.

Much of the neural parameter estimation literature uses global methods [6, 37, 99, 124]. Objections to local methods include their tendency to get stuck in local minima, and lack of insight on how model behavior varies over regions of parameter space. Sensitivities give some information on how firing patterns perturb, and can be used to calculate tangent directions for continuation (e.g. section 3.2). We have seen that different objective functions can improve the situation regarding local minima.

Several of the objective functions in the literature involve spike number [6, 124].

Spike number varies over the parameter ranges considered, resulting in discontinuities. Other objective functions involve smoothing [54] or transforming [116] the time series data. The objective functions we use are simple ones based directly upon biologically relevant quantities such as period and burst duration. This makes the progress of the optimization algorithms easier to interpret.

Optimization based upon dynamics. Consider bursting periodic data. Any reasonable fit to this data should both burst and be periodic. We use the geometry of bursting to introduce burst defining equations (Figure 2.1), and restrict to periodic solutions in our optimization iterates. The use of dynamical features such as periodic orbits and slow manifolds when applying optimization methods to ODEs was initiated by Casey [15] and is extended here to apply to bursting rhythms.

5.1 Future directions

The results of using our optimization methods on the Hindmarsh-Rose and pre-Botzinger complex models are promising, but there is much room for future work.

Application to different bursting models. The Hindmarsh-Rose and pre-Botzinger models are both square wave bursters. There are many other types of bursters, such as parabolic and elliptic bursters [4, 61]. Adapting our geometric methods to these other burster types would be useful. Spike number changes in both the Hindmarsh-Rose and preBotzinger models are associated with canards. Terman's [117] analysis on this topic considers systems where burst termination is associated with a fast subsystem homoclinic bifurcation. It would be interesting to investigate how spike number changes for bursters involving other types of bi-

furcations, and examine the effect of spike number on algorithm performance. We expect similar results to those in the Hindmarsh-Rose and preBotzinger systems.

The Hindmarsh-Rose and preBotzinger models each contain a single slow variable. This is the most common situation mathematically, but in reality multiple slow currents often occur. For example, Butera et al's [11] model ignores calcium currents which are known to exist in the preBotzinger complex. Optimization algorithms can aid in examining the effects of including additional currents. Casey [15] took this approach in comparing tonic spiking output from the Hodgkin-Huxley [59] and Clay [17] models. A similar comparison here between bursts from Butera et al's [11] model and one including calcium currents would be useful.

Multiple slow processes are considered in models of synaptically coupled preBotzinger cells [5]. Applying our burst event defining equations to networks of cells would be extremely useful, both for model calibration and for studying features such as phase response curves through event sensitivities. There is an interplay between synaptic inputs and intrinsic currents. The sensitivity calculations and optimization algorithms can be used to examine the roles of these different components in shaping network output. This is especially pertinent given that synaptic parameters can be difficult to measure experimentally.

Optimization algorithm refinements. There are many areas for further investigation and refinement in our optimization algorithms. Local minima are a common feature of ODE objective function landscapes [31]. These are particularly widespread for burst trajectories (section 3.2). Restricting to burst orbits and focusing on burst timing in our objective functions results in improved landscapes. Nevertheless, local minima still exist. These are associated with canards and chang-

ing spike number, which suggests looking for ways to change spike number appropriately when local minima are encountered. Developing automated methods to do this would be useful. The burst event and period sensitivities might be helpful in this regard. Canards are delicate and we expect burst timing to be very sensitive near a change in spike number.

The methods we use only consider solution trajectories to the ODE at each step. Curves which do not satisfy the ODE are not allowed. In addition we require that these solution trajectories be of a particular type, for example burst periodic orbits. Infinite barrier functions ensure that iterates remain in the valid domain. There are other approaches. For example, Bock [9] uses a multiple shooting method where separate solution trajectories are generated for each time interval. The initial conditions for each of these trajectory segments at the start of the optimization are chosen based upon the reference data. These segments are not initially required to match up at the end points. It is only by the end of the optimization iterates that a solution to the ODE is required.

Ramsay et al [102] introduce a method which uses basis functions (e.g. splines) to generate curves, together with penalty functions which measure the discrepancy between these curves and solutions to the ODE. The objective function is the sum of squared residuals of the basis function curve together with this penalty function. The basis function coefficients are computed in an inner optimization step. This inner optimization lies within an outer loop which treats the basis function coefficients as fixed and optimizes over the ODE parameters. Increasing the penalty function weight leads to closer approximations to solution curves of the ODE.

We have not addressed variance estimates for our optimization algorithms. Sta-

tistical inference for differential equations is largely unexplored [102] and there are many topics worth investigating. Incorporating the geometric ideas and numerical improvements from this thesis into an inference framework would be an important step.

BIBLIOGRAPHY

[1] J. Angstadt and W. Friesen. Synchronized oscillatory activity in leech neurons induced by calcium channel blockers. *Journal of Neurophysiology*, 66:1858–1873, 1991.

[2] M. Avoli, M. D'Antuono, J. Louvel, R. Kohling, G. Biagini, R. Pumain, G. D'Arcangelo, and V. Tancredi. Network and pharmacological mechanisms leading to epileptiform synchronization in the limbic system in vitro. *Progress in Neurobiology*, 68(3):167–207, 2002.

[3] T. Baker, D. Fuller, A. Zabka, and G. Mitchell. Respiratory plasticity: differential actions of continuous and episodic hypoxia and hypercapnia. *Respiration Physiology*, 129:25–35, 2001.

[4] R. Bertram, M. J. Butte, T. Kiemel, and A. Sherman. Topological and phenomenological classification of bursting oscillations. *Bulletin of Mathematical Biology*, 1995.

[5] J. Best, A. Borisyuk, J. Rubin, D. Terman, and M. Wechselberger. The dynamic range of bursting in a model respiratory pacemaker network. *SIAM Journal on Applied Dynamical Systems*, 4(4):1107–1139, 2005.

[6] U. Bhalla and J. Bower. Exploring parameter space in detailed single neuron models: simulations of the mitral and granule cells of the olfactory bulb. *Journal of Neurophysiology*, 69(6):1948–1965, 1993.

[7] U. Bickmeyer, M. Heine, T. Manzke, and D. Richter. Differential modulation of I_h by 5-HT receptors in mouse CA1 hippocampal neurons. *European Journal of Neuroscience*, 16:209–218, 2002.

[8] E. Bixler, A. Vgontzas, H. Lin, T. Ten Have, J. Rein, A. Vela-Bueno, and A. Kales. Prevalance of sleep-disordered breathing in women: effects of gender. *American Journal of Respiratory and Critical Care Medicine*, 163:608–613, 2001.

[9] H. Bock. Numerical treatment of inverse problems in chemical reaction kinetics. In *Modelling of chemical reaction systems*, Springer Series in Chemical Physics. Springer, 1981.

[10] A. Bonham. Neurotransmitters in the CNS control of breathing. *Respiration Physiology*, 101:219–230, 1995.

[11] R. J. Butera Jr., J. Rinzel, and J. C. Smith. Models of respiration rhythm generation in the pre-Botzinger complex. I. Bursting pacemaker neurons. *Journal of Neurophysiology*, 81:382–397, 1999.

[12] C. Canavier, J. Clark, and J. Byrne. Routes to chaos in a model of a bursting neuron. *Biophysical Journal*, 57(6):1245–1251, 1990.

[13] S. Cannon, R. Brown, and D. Corey. A sodium channel defect in hyperkalemic periodic paralysis: potassium-induced failure of inactivation. *Neuron*, 6:619–626, 1991.

[14] V. Carpentier, H. Vaudry, E. Mallet, A. Laquerriere, and P. Leroux. Increased density of somatostatin binding sites in respiratory nuclei of the brainstem in sudden infant death syndrome. *Neuroscience*, 86:159–166, 1998.

[15] R. Casey. *Periodic orbits in neural models: sensitivity analysis and algorithms for parameter estimation*. PhD thesis, Cornell University, 2004.

[16] T. R. Chay and J. Keizer. Minimal model for membrane oscillations in the pancreatic β-cell. *Biophysicla Journal*, 42:181–190, 1983.

[17] J. Clay. Excitability of the squid giant axon revisited. *Journal of Neurophysiology*, 80:903–913, 1998.

[18] W. Crill. Persistent sodium current in mammalian central neuons. *Annual Review of Physiology*, 58:349–362, 1996.

[19] G. S. Cymbalyuk and R. L. Calabrese. A model of slow plateau-like oscillations based upon the fast Na^+ current in a window mode. *Neurocomputing*, 38-40:159–166, 2001.

[20] O. Decroly and A. Goldbeter. From simple to complex oscillatory behavior - analysis of bursting in a multiply regulated biochemical system. *Journal of Theoretical Biology*, 124(2):219–250, 1987.

[21] R. Deisz, G. Fortin, and W. Zieglgansberger. Voltage dependence of excitatory postsynaptic potentials of rat neocortical neurons. *Journal of Neurophysiology*, 65:371–382, 1991.

[22] C. Del Negro and S. Chandler. Physiological and theoretical analysis of K^+ currents controlling discharge in neonatal rat mesencephalic trigeminal neurons. *Journal of Neurophysiology*, 77:537–553, 1997.

[23] R. Delgado-Lezama, J. Perrier, S. Nedergaard, G. Svirskis, and J. Hounsgaard. Metabotropic synaptic regulation of intrinsic response properties of spinal motoneurones. *Journal of Physiology (London)*, 504:97–102, 1997.

[24] J. Dennis and R. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM, Philadelphia, 1996.

[25] M. Diener. The canard unchained or how fast/slow dynamical systems bifurcate. *Mathematical Intelligencer*, 6(3):38–49, 1984.

[26] W. Eckhaus. A standard chase on French ducks. *Lecture notes in Mathematics*, 985:449–494, 1983.

[27] L. Edsberg and P. Wedin. Numerical tools for parameter-estimation in ODE systems. *Optimization methods and software*, 6(3):193–217, 1995.

[28] A. El Manira, J. Tegner, and S. Grillner. Calcium-dependent potassium channels play a critical role for burst termination in the locomotor network in lamprey. *Journal of Neurophysiology*, 72:1852–1861, 1994.

[29] F. Elsen and J. Ramirez. Calcium currents of rhythmic neurons recorded in the isolated respiratory network of neonatal mice. *Journal of Neuroscience*, 18:10652–10662, 1998.

[30] F. Elsen and J. Ramirez. Postnatal development differentially affects voltage-activated calcium currents in respiratory rhythmic versus nonrhythmic neurons of the pre-Botzinger complex. *Journal of Neurophysiology*, 94(2):1423–1431, 2005.

[31] W. Esposito and C. Floudas. Deterministic global optimization in nonlinear optimal control problems. *Journal of Global Optimization*, 17:97–126, 2000.

[32] J. Feldman and J. Smith. Cellular mechanisms underlying modulation of breathing pattern in mammals. *Annals of the New York Academy of Sciences*, 563:114–130, 1989.

[33] J. L. Feldman, G. S. Mitchell, and E. E. Nattie. Breathing: rhythmicity, plasticity, chemosensitivity. *Annual Review of Neuroscience*, 26:239–266, 2003.

[34] N. Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21(3):193–226, 1971.

[35] R. Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1961.

[36] R. Fletcher. *Practical methods of optimization*. John Wiley and Sons, 1987.

[37] W. Foster, L. Ungar, and J. Schwaber. Significance of conductances in Hodgkin-Huxley models. *Journal of Neurophysiology*, 70(6):2502–2518, 1993.

[38] R. Fregosi, Z. Luo, and M. Iizuka. GABA(A) receptors mediate postnatal depression of respiratory frequency by barbiturates. *Respiratory Physiology and Neurobiology*, 140(3):219–230, 2004.

[39] C. French and P. Gage. A threshold sodium current in pyramidal cells in rat hippocampus. *Neuroscience Letters*, 56:289–293, 1985.

[40] C. French, P. Sah, K. Buckett, and P. Gage. A voltage-dependent persistent sodium current in mammalian hippocampal neurons. *Journal of General Physiology*, 95:1139–1157, 1990.

[41] M. Fujii, K. Umezawa, and A. Arata. Dopamine desynchronizes the pacemaking neuronal activity of rat respiratory rhythm generation. *European Journal of Neuroscience*, 23(4):1015–1027, 2006.

[42] G. Funk, J. Smith, and J. Feldman. Generation and transmission of respiratory oscillations in medullary slices: role of excitatory amino acids. *Journal of Neurophysiology*, 70:1497–1515, 1993.

[43] R. Ghigliazza and P. Holmes. Minimal models of bursting neurons: how multiple currents, conductances, and timescales affect bifurcation diagrams. *SIAM Journal on Applied Dynamical Systems*, 3(4):636–670, 2004.

[44] M. S. Goldman, J. Golowasch, E. Marder, and L. F. Abbott. Global structure, robustness, and modulation of neuronal models. *Journal of Neuroscience*, 21(4):5229–5238, 2001.

[45] S. Gosgnach, G. Lanuza, S. Butt, H. Saueressig, Y. Zhang, T. Velasquez, D. Riethmacher, E. Callaway, O. Kiehn, and M. Goulding. V1 spinal neurons regulate the speed of vertebrate locomotor outputs. *Nature*, 440(7081):215–219, 2006.

[46] P. Gray, J. Rekling, C. Bocchiaro, and J. Feldman. Modulation of respiratory frequency by peptidergic input to rhythmogenic neurons in the preBotzinger complex. *Science*, 286:1566–1568, 1999.

[47] A. Griewank. *Evaluating derivatives: principles and techniques of algorithmic differentiation*. Frontiers in Applied Mathematics. SIAM: Philadelphia, 2000.

[48] A. Griewank, D. Juedes, and J. Utke. Algorithm 755: ADOL-C: a package for the automatic differentiation of algorithms written in C/C++. *ACM Transactions on Mathematical Software*, 22(2):131–167, 1996.

[49] J. Guckenheimer, K. Hoffman, and W. Weckesser. The forced van der Pol equation I: the slow flow and its bifurcations. *SIAM Journal on Applied Dynamical Systems*, 2(1):1–35, 2003.

[50] J. Guckenheimer and B. Meloon. Computing periodic orbits and their bifurcations with automatic differentiation. *SIAM Journal on Scientific Computing*, 22(3):951–985, 2000.

[51] J. Guckenheimer, J. Tien, and A. Willms. Bifurcations in the fast dynamics of neurons: implications for bursting. In S. Coombes and P. C. Bressloff, editors, *Bursting: the genesis of rhythm in the nervous system*. World Scientific Publishing Co., 2005.

[52] M. Haller, S. Mironov, A. Karschin, and D. Richter. Dynamic activation of K_{ATP} channels in rhythmically active neurons. *Journal of Physiology*, 537(1):69–81, 2001.

[53] R. Harris-Warrick, E. Marder, A. I. Selverston, and M. Moulins, editors. *Dynamic biological networks: the stomatogastric nervous system*. MIT Press, 1992.

[54] R. Hayes, J. Byrne, S. Cox, and D. Baxter. Estimation of single-neuron model parameters from spike train data. *Neurocomputing*, 65:517–529, 2005.

[55] B. Hille. *Ion channels of excitable membranes*. Sinauer, 3rd edition, 2001.

[56] J. Hindmarsh and P. Cornelius. The development of the hindmarsh-rose model for bursting. In S. Coombes and P. C. Bressloff, editors, *Bursting: the genesis of rhythm in the nervous system*, pages 3–18. World Scientific Publishing Co., 2005.

[57] J. Hindmarsh and R. Rose. A model of the nerve impulse using two first-order differential equations. *Nature, London*, 296:162–164, 1982.

[58] J. Hindmarsh and R. Rose. A model of neuronal bursting using 3 coupled 1st order differential equations. *Proceedings of the Royal Society of London Series B*, 221(1222):87–102, 1984.

[59] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology (London)*, 117:500–544, 1952.

[60] P. Holmes, R. Full, D. Koditschek, and J. Guckenheimer. The dynamics of legged locomotion: models, analyses, and challenges. *SIAM Review*, 48(2):207–304, 2006.

[61] F. C. Hoppensteadt and E. M. Izhikevich. *Weakly connected neural networks*. Springer-Verlag, 1997.

[62] K. Huey, I. Brown, M. Jordan, and F. Powell. Changes in dopamine D(2)-receptor modulation of the hypoxic ventilatory response with chronic hypoxia. *Respiration Physiology*, 123:177–187, 2000.

[63] K. Huey, M. Low, M. Kelly, R. Juarez, J. Szewczak, and F. Powell. Ventilatory responses to acute and chronic hypoxia in mice: effects of dopamine D(2)-receptors. *Journal of Applied Physiology*, 89:1142–1150, 2000.

[64] E. Izhikevich. Neural excitability, spiking and bursting. *International Journal of Bifurcation and Chaos*, 10(6):1171–1266, 2000.

[65] S. Johnson, J. Smith, and J. Feldman. Modulation of $G_{i/o}$ protein-mediated mechanisms. *Journal of Applied Physiology*, 80:2120–2133, 1996.

[66] S. Johnson, J. Smith, G. Funk, and J. Feldman. Pacemaker behavior of respiratory neurons in medullary slices from neonatal rat. *Journal of Neurophysiology*, 72:2598–2608, 1994.

[67] A. Kahn, D. Blum, E. Rebuffat, M. Sottiaux, J. Levitt, A. Bochner, M. Alexander, J. Grosswasser, and M. Muller. Polysomnographic studies of infants who subsequentyl died of sudden infant death syndrome. *Pediatrics*, 82(5):721–727, 1988.

[68] J. Keener and J. Sneyd. *Mathematical physiology*. Springer-Verlag, 1998.

[69] O. Kiehn, O. Kjaerulff, M. Tresch, and R. Harris-Warrick. Contributions of intrinsic motor neuron properties to the production of rhythmic motor output in the mammalian spinal cord. *Brain Research Bulletin*, 53(5):649–659, 2000.

[70] R. Kinkead, K. Bach, S. Johnson, B. Hodgeman, and G. Mitchell. Plasticity in respiratory motor control: intermittent hypoxia and hypercapnia activate opposing serotonergic and noradrenergic modulatory systems. *Comparative Biochemistry and Physiology A. Molecular and Integrative Physiology*, 130:207–218, 2001.

[71] H. Kinney, J. Filiano, and W. White. Medullary serotonergic network deficiency in the sudden infant death syndrome: review of a 15-year study of a single data set. *Journal of Neuropathology and Experimental Neurology*, 60:228–247, 2001.

[72] S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[73] N. Koshiya and J. C. Smith. Neuronal pacemaker for breathing visualized in vitro. *Nature*, 1999.

[74] C. K.R.T. Jones. Geometric singular perturbation theory. In L. Arnold, editor, *Dynamical Systems, Montecatini Terme*, volume 1609 of *Lecture Notes in Mathematics*, pages 44–118. Springer-Verlag, Berlin, 1994.

[75] Y. A. Kuznetsov. *Elements of applied bifurcation theory*. Springer, 2nd edition, 1998.

[76] S. Lahiri and R. Fitzgerald. Reflex responses to chemoreceptor stimulation. In *Handbook of physiology. The Respiratory System. Control of breathing*, volume 2, chapter 10, pages 313–362. American Physiological Society, 1986.

[77] P. Lalley. μ-Opioid receptor agonist effect on medullary respiratory neurons on the cat: evidence for involvement in certain types of ventilatory disturbances. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 285:R1287–R1304, 2003.

[78] K. Levenberg. A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.*, 2:164–168, 1944.

[79] A. Lijowska, N. Reed, B. Chiodini, and B. Thachy. Sequential arousal and airway-defensive behavior of infants in asphyxial sleep environments. *Journal of Applied Physiology*, 83:219–228, 1997.

[80] I. Llona and J. Eugenin. Central actions of somatostatin in the generation and control of breathing. *Biological Research*, 38:347–352, 2005.

[81] J. Magistretti and A. Alonso. Biophysical properties and slow voltage-dependent inactivation of a sustained sodium current in entorhinal cortex layer-II principal neurons. *Journal of General Physiology*, 114:491–509, 1999.

[82] E. Marder and R. Calabrese. Principles of rhythmic motor pattern generation. *Physiological Reviews*, 76(3):687–717, 1996.

[83] D. Marquardt. An algorithm for least squares estimation of nonlinear parameters. *SIAM Journal*, 11:431–441, 1963.

[84] R. Mellins, H. Balfour Jr., G. Turino, and R. Winters. Failure of automatic control of ventilation (Ondine's curse): report of an infant born with this syndrome and review of the literature. *Medicine (Baltimore)*, 49:487–504, 1970.

[85] S. Mironov, N. Hartelt, and M. Ivannikov. Mitochondrial K_{ATP} channels in respiratory neurons and their role in hypoxic facilitation of rhythmic activity. *Brain Research*, 1033:20–27, 2005.

[86] S. Mironov and D. Richter. Intracellular signalling pathways modulate K_{ATP} channels in inspiratory brainstem neurones and their hypoxic activation: involvement of metabotropic receptors, G-proteins and cytoskeleton. *Brain Research*, 853:60–67, 2000.

[87] G. Mitchell and S. Johnson. Neuroplasticity in respiratory motor control. *Journal of Applied Physiology*, 94:358–374, 2003.

[88] M. Mitchell. *An introduction to genetic algorithms*. MIT Press, 1996.

[89] J. More. The Levenberg-Marquardt algorithm: implementation and theory. In G. Watson, editor, *Numerical analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 105–116. Springer-Verlag, Berlin, 1977.

[90] E. Nattie. CO_2, brainstem chemoreceptors and breathing. *Progress in Neurobiology*, 59:299–331, 1999.

[91] J. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.

[92] H. Onimaru, A. Arata, and I. Homma. Firing properties of respiratory rhythm generating neurons in the absence of synaptic transmission in rat medulla in vitro. *Experimental Brain Research*, 76:530–536, 1989.

[93] V. Osipov and E. Ponizovskaya. Stochastic resonance in the Brusselator model. *Physical Review E*, 61(4):4603–4605, 2000.

[94] P. Pahapill and A. Lozano. The pedunculopontine nucleus and Parkinson's disease. *Brain*, 123:1767–1783, 2000.

[95] O. Paulsen and T. Sejnowski. Natural patterns of activity and long-term synaptic plasticity. *Current Opinion in Neurobiology*, 10(2):172–179, 2000.

[96] E. T. Phipps. *Taylor series integration of differential-algebraic equations: automatic differentiation as a tool for simulating rigid body mechanical systems*. PhD thesis, Cornell University, 2003.

[97] R. Plant. Bifurcation and resonance in a mathematical model for bursting nerve cells. *Journal of Mathematical Biology*, 11:15–32, 1981.

[98] C. Poets, R. Meny, M. Chobanian, and R. Bonofiglo. Gasping and other cardiorespiratory patterns during sudden infant deaths. *Pediatric Research*, 45:350–354, 1999.

[99] A. A. Prinz, C. P. Billimoria, and E. Marder. Alternative to hand-tuning conductance-based models: construction and analysis of databases of model neurons. *Journal of Neurophysiology*, 90:3998–4015, 2003.

[100] M. Rabinovich and H. Abarbanel. The role of chaos in neural systems. *Neuroscience*, 87(1):5–14, 1998.

[101] J. Ramirez, A. Tryba, and F. Pena. Pacemaker neurons and neuronal networks: an integrative view. *Current Opinion in Neurobiology*, 14(6):665–674, 2004.

[102] J. Ramsay, G. Hooker, J. Cao, and D. Campbell. Estimating differential equations. Submitted.

[103] D. Richter, T. Manzke, B. Wilken, and E. Ponimaskin. Serotonin receptors: guardians of stable breathing. *Trends in Molecular Medicine*, 9(12):542–548, 2003.

[104] J. Rinzel and Y. S. Lee. Dissection of a model for neuronal parabolic bursting. *Journal of Mathematical Biology*, 25:653–675, 1987.

[105] P. Roper, C. H. Brown, C. W. Bourque, and W. E. Armstrong. Autoregulation of bursting of AVP neurons of the rat hypothalamus. In S. Coombes and P. Bresloff, editors, *Bursting: the genesis of rhythm in the nervous system*, pages 49–88. World Scientific Publishing Co., 2005.

[106] H. Sabbagh. Control of chaotic solutions of the Hindmarsh-Rose equations. *Chaos Solitons and Fractals*, 11(8):1213–1218, 2000.

[107] T. Santiago and N. Edelman. Opioids and breathing. *Journal of Applied Physiology*, 59:1675–1685, 1985.

[108] V. Schechtman, R. Harper, A. Wilson, and D. Southall. Sleep apnea in infants who succumb to the sudden infant death syndrome. *Pediatrics*, 87:841–846, 1991.

[109] K. Schittkowski. Parameter estimation in differential equations. In R. Agarwal, editor, *Recent trends in optimization theory and applications*, volume 5 of *World Scientific Series in Applicable Analysis*, pages 353–370. World Scientific Publishing Co., 1995.

[110] M. Segal. Endogenous bursts underlie seizurelike activity in solitary excitatory hippocampal neurons in microcultures. *Journal of Neurophysiology*, 72:1874–1884, 1994.

[111] A. Sherman. Contributions of modeling to understanding stimulus-secretion coupling in pancreatic β cells. *American Journal of Physiology - Endocrinology and Metabolism*, 271(2):E362–E372, 1996.

[112] A. Sherman and J. Rinzel. Rhythmogenic effects of weak electrotonic coupling in neuronal models. *Proceedings of the National Academy of Sciences USA*, 89:2471–2474, 1992.

[113] A. Sherman, J. Rinzel, and J. Keizer. Emergence of organized bursting in clusters of pancreatic beta cells by channel sharing. *Biophysical Journal*, 54:411–425, 1988.

[114] J. Smith, H. Ellenberger, K. Ballanyi, D. Richter, and J. Feldman. Pre-Botzinger complex: a brainstem region that may generate respiratory rhythm in mammals. *Science*, 254:726–729, 1991.

[115] S. Stojilkovic, H. Zemkova, and F. van Goor. Biophysical basis of pituitary cell type-specific Ca^{2+} signaling-secretion coupling. *Trends in Endocrinology and Metabolism*, 16(4):152–159, 2005.

[116] J. Tabak, C. Murphey, and L. Moore. Parameter estimation methods for single neuron models. *Journal of Computational Neuroscience*, 9(3):215–236, 2000.

[117] D. Terman. Chaotic spikes arising from a model of bursting in excitable membranes. *SIAM Journal on Applied Mathematics*, 51(5):1418–1450, 1991.

[118] M. Thoby-Brisson, B. Cauli, J. Champagnat, G. Fortin, and D. Katz. Expression of functional tyrosine kinase B receptors by rhythmically active respiratory neurons in the preBotzinger complex of neonatal mice. *Journal of Neuroscience*, 23(20):7685–7689, 2003.

[119] M. Thoby-Brisson and J.-M. Ramirez. Identification of two types of inspiratory pacemaker neurons in the isolated respiratory neural network of mice. *Journal of Neurophysiology*, 2001.

[120] J. H. Tien, W. D. Hazelton, R. Sparks, and C. M. Ulrich. A Michaelis-Menten-style model for the autocatalytic enzyme prostaglandin H synthase. *Bulletin of Mathematical Biology*, 67:683–700, 2005.

[121] A. Tryba, F. Pena, and J. Ramirez. Gasping activity in vitro: a rhythm dependent on 5-HT2A receptors. *Journal of Neuroscience*, 26(10):2623–2634, 2006.

[122] A. Tryba and J. Ramirez. Hyperthermia modulates respiratory pacemaker bursting properties. *Journal of Neurophysiology*, 92(5):2844–2852, 2004.

[123] B. van der Pol. On "relaxation oscillations". *Philosophical Magazine*, 2(11):978–992, 1926.

[124] M. C. Vanier and J. M. Bower. A comparative survey of automated parameter-search methods for compartmental neural models. *Journal of Computational Neuroscience*, 7:149–171, 1999.

[125] J. Viemari and J. Ramirez. Norepinephrine differentially modulates different types of respiratory pacemaker and nonpacemaker neurons. *Journal of Neurophysiology*, 95(4):2070–2082, 2006.

[126] M. Wall and N. Dale. A slowly activating Ca^{2+}-dependent K^+ current that plays a role in termination of swimming in Xenopus embryos. *Journal of Physiology (London)*, 487:557–572, 1995.

[127] H. Wang, R. Stornetta, D. Rosin, and P. Guyenet. Neurokinin-1 receptor-immunoreactive neurons of the ventral respiratory group in the rat. *Journal of Comparative Neurology*, 434(2):128–146, 2001.

[128] X. J. Wang. Multiple dynamical modes of thalamic relay neurons: rhythmic bursting and intermittent phase-locking. *Neuroscience*, 59:21–31, 1994.

[129] C. Washburn, D. Bayliss, and P. Guyenet. Cardiorespiratory neurons of the rat ventrolateral medulla contain TASK-1 and TASK-3 channel mRNA. *Respiratory Physiology and Neurobiology*, 138:19–35, 2003.

[130] C. Washburn, J. Sirois, E. Talley, P. Guyenet, and D. Bayliss. Serotonergic raphe neurons express TASK channel transcripts and a TASK-like pH-and halothane-sensitive K^+ conductance. *Journal of Neuroscience*, 22(4):1256–1265, 2002.

[131] Y. Yamamoto, H. Onimaru, and I. Homma. Effect of substance P on respiratory rhythm and pre-inspiratory neurons in the ventrolateral structure of rostral medulla oblongata: an in vitro study. *Brain Research*, 599:272–276, 1992.